

AD-A031 256

SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF ELECTRICAL--ETC F/G 9/4
DETECTION AND ESTIMATION OF SIGNALS IN UNCERTAIN ENVIRONMENTS.(U)
AUG 76 M D SRINATH

UNCLASSIFIED

AFOSR-TR-76-1119

NL

1 OF 1
AD
A031256



END

DATE
FILMED
11-1-76

AD A031256

AFOSR - TR - 76 - 1119

(Handwritten: K)

Final Report

DETECTION AND ESTIMATION OF SIGNALS
IN UNCERTAIN ENVIRONMENTS

See Form 14/1/82

Electrical Engineering Department

Institute of Technology

Southern Methodist University

Dallas, Texas 75275

DDC
RECEIVED
OCT 27 1976
(Handwritten: J)

Approved for public release;
distribution unlimited.



DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

DETECTION AND ESTIMATION OF SIGNALS
IN UNCERTAIN ENVIRONMENTS

Final Report

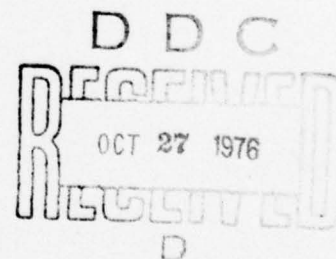
Prepared for

Directorate of Mathematical and Information Sciences
Air Force Office of Scientific Research (AFSC)
Bolling Air Force Base, D.C. 20332

Under Grant No. AFOSR 74-2727

by

M. D. Srinath
Principal Investigator
Southern Methodist University
Dallas, Texas 75275



31st Aug 1976

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

Abstract

Recent papers have brought out the intimate connection that exists between the problems of detection and estimation in many communication problems. In many cases, it becomes necessary to simultaneously detect and extract the signal so that the two operations must be jointly optimized in order to improve receiver performance. However, the receiver structures so obtained are considerably more complicated than conventional receivers and it is not clear that this added complexity results in a significant improvement in receiver performance. This report investigates a specific suboptimal scheme for estimating a signal in the presence of uncertainty regarding its reception, namely, the probabilistic estimation scheme. The scheme uses a probabilistic judgement to determine the presence of the signal in the observations. In order to obtain explicit estimation algorithms and to facilitate performance evaluation, a specific problem of signal estimation under uncertainty, namely the tracking of a target in a multi-target multi-sensor environment is considered. Explicit algorithms for this problem are developed and the performance of the receiver is evaluated under conditions involving fading media or quantized observations.

RECEIVED BY	
RTD	White Section <input checked="" type="checkbox"/>
DD	Buff Section <input type="checkbox"/>
CHAS. CHURCH	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISSEMINATION/REPRODUCTION CODES	
FILE	CLASS. AND OF SPECIAL
A	

TABLE OF CONTENTS

1. Introduction
2. Signal and Observation Models
 - 2.1 Introduction
 - 2.2 Error Bounds
3. The Probabilistic Estimator
 - 3.1 Introduction
 - 3.2 Probabilistically Supervised Estimator
 - 3.3 Evaluation of Probabilistic Estimator
 - 3.3.1 Suboptimal Estimator of Finite Memory
 - 3.3.2 Suboptimal Estimator with Gating
 - 3.4 Simulations
4. Coupled Risk Formulation
 - 4.1 Introduction
 - 4.2 Coupled Risk Formulation
 - 4.3 Suboptimal Estimator and Detector
 - 4.4 Simulation Results
5. Tracking Over Fading Channels
6. Estimation Using Quantized Measurements
 - 6.1 Introduction
 - 6.2 Probabilistically Supervised Estimation with Quantized Measurements
 - 6.3 Simulations
7. Conclusions
- Appendix A: Statistics of Extraneous Returns
- Appendix B: Target Tracking Over Fading Channels

1. INTRODUCTION

Problems of signal estimation and detection which are central to communication systems have received considerable attention in recent years. While for the most part the solutions to these problems have been obtained assuming that they are essentially independent operations [1], recent papers have brought out the intimate connection that exists between the two [2-9]. In many cases, it becomes necessary to simultaneously detect and extract the signal so that the two operations must be coupled in order to improve receiver performance. Such examples occur in all circumstances where properties of the signal source or transmission medium, signal presence or absence, or signal identification are jointly desired and may arise in such diverse fields as underwater communications, seismic detection, pattern recognition or radar ranging and tracking. For example, in communication over fading dispersive channels, the exact value of channel attenuation is not known. In a multitarget, multisensor environment, it becomes necessary to simultaneously identify and track the target. In a pattern classification problem, the particular pattern giving rise to the received signal is unknown.

The basic problem may be formulated as follows. The data z received during each observation interval I consists of (a) a signal y corrupted by observation noise v (Hypothesis H_1) or (b) noise v alone (Hypothesis H_0). We assume that z and y are vectors consisting of sampled values. The signal y is assumed to be a realization of a stochastic process. We are required to recursively estimate either the signal y or some parameters θ associated with it.

We thus write

$$\begin{aligned} H_1 : z_k &= y_k + v_k \\ H_0 : z_k &= v_k \end{aligned} \quad \text{for } k \in I \quad (1.1)$$

The receiver is required to make two functionally related decisions at the end of each observation interval, the first on the presence or absence of the signal (a detection decision) and the second on the value of the signal (or parameter) vector (an estimation decision).

Let us define R_D as the risk associated with detecting the signal presence and R_E as the risk associated with estimating the signal. If the two operations of detection and estimation can be carried out in parallel or simultaneously, but independently and the two components R_D and R_E can be minimized separately, then the two problems are said to be uncoupled [4,5]. The detection procedure is carried out in the usual fashion but the estimator performance has to be optimized in the face of uncertainty as to the signal presence. In many situations such a separation of costs is not possible and the two operations are said to be coupled.

We can write the observation equation for the joint detection-estimation problem as

$$z_k = \beta y_k + v_k \quad \text{for } k \in I \quad (1.2)$$

where β is a binary random variable taking on the values 0 or 1. The problem is then referred to as a single-shot problem, and the same hypothesis is true over the entire observation interval. β is called the indicator variable.

Let Z_k denote the sequence $[z_1, z_2, \dots, z_k]$. Assuming a quadratic cost for the estimation problem, let the estimate of the signal y_k be

denoted as \hat{y}_k and let \hat{y}_k^1 denote the estimate assuming hypothesis H_1 to be true. It is clear that $\hat{y}_k^0 = E\{y_k\} = \mu_{y_k}$.

By expanding in mixtures, we can then write the estimate as

$$\hat{y}_k = \sum_{i=0,1} y_k^i P(H^i | Z_k) \quad (1.3)$$

If P_0 and P_1 are the prior probabilities associated with H_0 and H_1 , define

$$\Lambda_k = \frac{P_1}{P_0} \frac{p(Z_k | H_1)}{p(Z_k | H_0)} \quad (1.4)$$

and

$$\Gamma_k = \frac{\Lambda_k}{1 + \Lambda_k} \quad (1.5)$$

Then an alternative form of Eq. (1.3) is [8]

$$y_k = y_k^1 + [1 - \Gamma_k] \mu_{y_k} \quad (1.6)$$

This differs from the form obtained by Middleton and Esposito [4] in the presence of the second term.

It can be further shown [11] that

$$\Gamma_k = E\{\beta | Z_k\} \quad (1.7)$$

It can be seen from Eq. (1.4) that obtaining the optimal estimate requires computation of the likelihood function Λ_k . The structure of the optimal estimator consists of an estimator for each hypothesis, the outputs of which are combined in a suitably weighted fashion to give the optimal estimate.

If the active hypothesis changes from observation to observation, we have the multi-shot problem. In this case, β_k can be modeled as an independent random sequence taking on the values 0 and 1. At stage k , the possible values of $\beta_1, \beta_2 \dots \beta_k$ come from a set of 2^k binary sequences

of length k . Let these sequences be denoted by σ_j^k , $j = 1, 2 \dots k$.

Then we can write

$$\hat{y}_k = \sum_{j=1}^{2^k} E\{y_k | z_k, \sigma_j^k\} p\{\sigma_j^k | z_k\} \quad (1.8)$$

The estimator at the k^{th} stage consists of a bank of 2^k filters (one for each possible sequence), whose outputs are then combined appropriately.

It is clear from Eq. (1.8) that this leads to an exploding memory requirement.

It is well-known [12] that the exploding memory problem can be replaced by one that requires storing complete density functions. The relevant densities may be recursively computed if there exists a fixed dimensional sufficient statistic.

Various suboptimal schemes have been proposed to eliminate the exploding memory requirements of Eq. (1.8). The linear estimator [13] does provide one such scheme in which a single filter is used to provide the best linear estimator. Other schemes essentially choose one or a finite number of the 2^k possible sequences σ_j^k in some manner. This reduces the summation in Eq. (1.7) to a summation over a fixed number of terms and hence leads to a fixed number of estimators in the implementation.

In this report, we will investigate a specific suboptimal scheme for estimating a signal in the presence of uncertainty regarding its reception, namely, the probabilistic estimation scheme [14,15]. Explicit algorithms for estimation using the scheme are developed and the performance of the estimator under several conditions is evaluated. In order to obtain explicit algorithms, a specific problem of signal estimation under uncertainty, namely the tracking of a target in a multi-target

multi-sensor environment is considered. While the algorithms presented in this report have been developed explicitly for this problem, they can easily be extended to other situations involving similar models.

The report is organized as follows. In Chapter II, the specific problem considered in the report is formulated and the observation models presented. The structure of the probabilistic estimator for this problem is developed in Chapter III and its performance compared with that of two other suboptimal estimators presented earlier in the literature.

As indicated earlier, in all problems where there is uncertainty regarding the signal presence in the observations, the two operations of signal detection and estimation must be carried out simultaneously. Thus the cost associated with these two problems must be optimized jointly. However the receiver structures so obtained are usually complicated. Hence it is desirable to compare their performance with suboptimal receivers to determine whether a significant improvement results from the increased complexity of the receiver. Chapter IV presents such a comparison between a receiver obtained using the coupled risk formulation and the scheme. Chapter V presents an evaluation of the probabilistic scheme in fading media while Chapter VI develops algorithms for estimation with quantized observations.

REFERENCES

- [1] A. P. Sage and J. L. Melsa, ESTIMATION THEORY WITH APPLICATIONS TO COMMUNICATION AND CONTROL, McGraw-Hill Book Co., New York, 1971.
- [2] T. Kailath, "A Note on Least Square Estimates from Likelihood Ratios," INFORMATION AND CONTROL, Vol. 13, Dec. 1968, pp. 534-540.
- [3] R. Esposito, "On a Relation between Detection and Estimation in Decision Theory, " INFORMATION AND CONTROL, Vol. 12, Feb. 1968, pp. 116-120.
- [4] D. Middleton and R. Esposito, "Simultaneous Optimum Detection and Estimation of Signals in Noise," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-14, May 1968, pp. 434-444.
- [5] D. Middleton and R. Esposito, "New Results in the Theory of Simultaneously Optimum Detection and Estimation of Signals in Noise," PROBLEMY PEREDACHI INFORMATSII, Vol. 6, 1970, pp. 3-20.
- [6] R. L. Stratanovich, "Detection and Estimation of Signals in Noise when One or Both Are Non-Gaussian," PROC. IEEE, Vol. 58, May 1970, pp. 670-679.
- [7] M. D. Srinath and P. K. Rajasekaran, "Estimation of Randomly Occuring Stochastic Signals in Gaussian Noise," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-17, March 1971, p. 206.
- [8] A. G. Jaffer and S. C. Gupta, "On Relations Between Detection and Estimation of Discrete-Time Processes," INFORMATION AND CONTROL, Vol. 20, No. 1, Feb. 1972, pp. 46-54.
- [9] T. G. Birdsall and J. O. Gobien, "Sufficient Statistics and Reproducing Densities in Simultaneous Sequential Detection and Estimation", IEEE TRANS. ON INFORMATION THEORY, Vol. IT-19, No. 6, Nov. 1973, pp. 760-768.
- [10] A. Fredriksen, D. Middleton, and D. Vandelinde, "Simultaneous Signal Detection and Estimation under Multiple Hypothesis," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-18, No. 5, Sept. 1972, pp. 607-613.
- [11] D. G. Lainiotis, "On a General Relationship between Estimation, Detection and the Bhattacharya Coefficient," IEEE TRANS. ON INFORMATION THEORY, Vol IT-15, 1969, pp. 504-505.
- [12] S. C. Fralick, "Learning to Recognize Patterns without a Teacher," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-13, Jan. 1967, pp. 57-64.

- [13] N. E. Nahi, "Optimal Recursive Estimation with Uncertain Observation," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-14, May 1968, p. 434.
- [14] A. K. Agrawala, "Learning with a Probabilistic Teacher," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-16, July 1970, pp. 373-379.
- [15] P. K. Rajasekaran and M. D. Srinath, "Adaptive Estimation with Uncertain Observation," PROCEEDINGS OF THE FIFTH ASILOMAR CONFERENCE ON CIRCUITS AND SYSTEMS, California, Nov. 1972.
- [16] A. G. Jaffer and S. C. Gupta, "Coupled Detection-Estimation of Gaussian Processes in Gaussian Noise," IEEE TRANS. ON INFORMATION THEORY, Vol. IT-18, No. 1, Jan. 1972, pp. 106-110.

2. SIGNAL AND OBSERVATION MODELS

2.1 Introduction

The problem of tracking a target such as a space vehicle in multi-target and multi-sensor environments has received considerable attention recently [1-3]. When observations from several sensors are supplied to the receiver, it must decide which, if any, of the sensor measurements corresponds to the target being tracked and suitably process these measurements to update the estimate of the target state. In real time surveillance systems, tracking filter performance heavily depends on two functions. The first function is referred to as sensor return-to-track correlation [3-5]. In the presence of multiple targets a sensor return could originate from one of three possible sources (hypotheses). The return could be from the object being tracked, could originate from one of the other objects not being tracked, or could be noise alone. Sensor return-to-track correlation is a decision process serving to decide whether a sensor return originated from the object being tracked. The second function, tracking filtering, has been developed in [5]. The sensor return which originates from the object being tracked is used in track updating; otherwise, there is no updating and the one-stage prediction is used. The dynamic equation of motion of the target being tracked is assumed to be modelled by a set of Gauss-Markov difference equations

$$\begin{aligned}x_{k+1} &= \phi_k x_k + \Gamma_k \omega_k \\y_k &= c_k x_k\end{aligned}\tag{2.1}$$

where x_k is the $n \times 1$ state vector, and y_k is a $p \times 1$ output vector. The m -vector ω_k is white zero mean gaussian with covariance $E\{\omega_k \omega_k^T\} = Q_k \delta_{ke} \phi_k$ is an $n \times n$ state transition matrix and Γ_k is an $n \times m$ excitation matrix.

If a particular sensor is not locked onto the target being tracked, the corresponding sensor return at any stage can originate from one of three sources (hypotheses), namely, (i) noise alone (e.g., cut-off communication link), (ii) false alarm (e.g. thermal or process noise, or clutter from a vehicle not being tracked), and (iii) the vehicle being tracked. The receiver performs under uncertainty of reception in a ternary case which considers all the possibilities. A suitable model for the observations under each hypothesis can then be formulated as follows [1]:

Under H^1 , the observation z_k is given by

$$H_k^1 : z_k = u_k \quad (2.2)$$

where the $p \times 1$ vector u_k is a zero mean, white Gaussian sequence with covariance $E\{u_k u_k^T\} = R_k \delta_{km}$. Under hypothesis H^2 , the current observation is not used in track updating and the one-stage prediction is used instead. The effect of the extraneous returns can be modeled as a white, zero mean sequence v_k , with covariance $E\{v_k v_k^T\} = \Omega_k \delta_{km}$. It is therefore reasonable to model the observation z_k under hypothesis H^2 as

$$H_k^2 : z_k = c_k \hat{x}(k|k-1) + v_k \quad (2.3)$$

where $\hat{x}(k|k-1)$ is the one-stage prediction of x_k .

The expression for the covariance kernel, Ω_k , is given in [5] and for $p = 2$ simplifies to

$$\Omega_k = \frac{C_k^T V_{k|k-1} C_k + R_k}{1 + \frac{\pi}{2} \cdot n_{TF} \sqrt{1-\rho^2}} \quad (2.4)$$

where $V_{k|k-1} = E\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\}$, n_{TF} is the expected number of incorrect returns, in the one sigma region

$\sigma_k = \min_i [(C_k^T V_{k|k-1} C_k + R_k)_{ii}]^{1/2}$ and ρ is the correlation between two measurements.

The noise sequences ω_k , v_k , and u_k are mutually independent and are uncorrelated with respect to the initial state x_0 which is multivariate normally distributed of dimension n with mean μ_0 and covariance V_0 , $N_n[\mu_0, V_0]$.

Under H^3 , the observation corresponds to the vehicle being tracked and hence we have

$$H_k^3 : z_k = c_k x_k + u_k$$

It is well known that the optimal Bayes estimator for this problem involves an evergrowing memory and is hence not feasible. This is because the estimate of x_k at each stage depends on the sequence of past true hypotheses. This sequence, however, is unknown and hence the estimate must be averaged over all possible past sequences. This corresponds to the unsupervised learning problem in pattern classification. It is clear that as the number of stages increases, the total number of sequences over which the estimates must be averaged increases very rapidly. We, therefore, need to resort to suboptimal schemes which use a fixed memory.

2.2 Error Bounds

As indicated above, the case in which the hypothesis at each stage associated with any sensor return is not known, corresponds to unsupervised learning. If the hypotheses are known, we have supervised learning. In general, if the classifications of the sensors are not known a priori, unsupervised learning must be resorted to. In many cases partial information about the classifications can be gained by making use of all the past observations as well as all the past classifications; e.g., by probabilistic judgements to determine the classifications; this is referred to as partially supervised learning, Ξ_k^p .

Intuitively, the variance associated with partially supervised learning is bounded below by that of supervised learning (learning with a teacher), and bounded above by that of unsupervised learning (optimal Bayesian learning). These bounds for the optimal stochastic control problem with finite state space have been investigated in [6], and tighter bounds given in [7]. Similar bounds for the optimal filtering problem can be established by augmenting the message model of Eq. (1) such that

$$x_{k+1}^1 = \phi_k x_k^1 + \Gamma_k \omega_k^1 \quad (2.6)$$

$$x_{k+1}^2 = f(x_k^2, \omega_k^2), \quad x_k^2 \in \{0,1\} \quad \forall k \quad (2.7)$$

where Eq. (2.6) can be identified with Eq. (2.1) by setting $x_k^1 = x_k$ and $\omega_k^1 = \omega_k$, Eq. (2.7) is the model for switching hypotheses, ω_k^2 is white zero mean gaussian sequence uncorrelated with ω_k^1 as well as the initial state, and x_k^1 and x_k^2 are independent.

The corresponding observation model is

$$z_k = h(x_k^1, x_k^2) + \epsilon_k \quad (2.8)$$

which can be identified with Eqs. (2.2) - (2.4) by setting

$$\begin{aligned} x_k^1 = 0, \quad x_k^2 = 0, \quad \text{and } \epsilon_k = v_k, \quad \text{i.e. } h[x_k^1, x_k^2] = 0 \\ \text{if } H_k^1 \text{ is active} \end{aligned} \quad (2.9)$$

$$\begin{aligned} x_k^1 = 0, \quad x_k^2 = 1, \quad \text{and } \epsilon_k = v_k, \\ \text{i.e. } h[x_k^1, x_k^2] = x_k^2 C_k x_{k|k-1}^1 + v_k \quad \text{if } H_k^2 \text{ is active} \end{aligned} \quad (2.10)$$

and

$$\begin{aligned} x_k^1 = x_k, \quad x_k^2 = 1 \quad \text{and } \epsilon_k = v_k, \\ \text{i.e. } h[x_k^1, x_k^2] = x_k^2 C_k x_k^1 + v_k \quad \text{if } H_k^3 \text{ is active} \end{aligned} \quad (2.11)$$

Let $V_R^S(\omega_k)$, $V_R^U(\omega_k)$ and $V_k^P(\omega_k)$ denote the error variances corresponding to the supervised estimator (optimal Bayesian estimator with known classification of sensor returns), the unsupervised estimator (optimal Bayesian estimator with unclassified sensor returns) and partially supervised estimator. Then, following an argument similar to [10], we can show that

$$V_k^S(\omega(k)) \leq V_k^P(\omega(k)) \leq V_k^U(\omega(k)) \quad \forall k \quad (2.12)$$

That is, the error variance associated with the partially supervised estimator is bounded below by that of the optimal Bayesian estimator with known classifications and above by the estimator with unknown classifications.

The proof of Eq. (2.12) is by induction and is as follows.

At the final stage N , the optimal estimate \hat{x}_N and the corresponding optimal return $V_N(\omega(N))$ are evaluated, respectively, as

$$\hat{x}_N = \text{conditional mean of } f(x_N | Z_N) \quad (2.13)$$

and

$$V_N(\omega(N)) = \text{conditional variance of } f(x_N | Z_N) \quad (2.14)$$

where $\omega(N) = f(x_N | Z_N)$

Hence, at the final stage N

$$\begin{aligned} \min_{\hat{x}_N} E C(\hat{x}_N, x_N, N) &= E \min_{Z_N} E_{\hat{x}_N} C(\hat{x}_N, x_N, N) \\ &= E_{Z_N} V_N(\omega(N)) \end{aligned} \quad (2.15)$$

where E_{Z_N} denotes the mathematical expectation with respect to the distribution of Z_N , and $E_{\hat{x}_N}$ denotes the mathematical expectation with respect to the conditional distribution of x_N given Z_N .

Assume that the following equality holds at stage k+1

$$\min_{\hat{x}_{k+1} \dots \hat{x}_N} E \sum_{i=k+1}^N C(\hat{x}_i, x_i, i) = E_{Z_{k+1}} V_{k+1}(\omega(k+1)) \quad (2.16)$$

Then, at stage k, application of the principle of optimality yields

$$\begin{aligned} \min_{\hat{x}_k \dots \hat{x}_N} E \sum_{i=1}^N C(\hat{x}_i, x_i, i) &= \min_{\hat{x}_k} \left\{ E C(\hat{x}_k, x_k, k) + E_{Z_{k+1}} V_{k+1}(\omega(k+1)) \right\} \\ &= \min_{\hat{x}_k} \left\{ E_{Z_k} E_{Z_{k+1} | Z_k} C(\hat{x}_k, x_k, k) + E_{Z_k} E_{Z_{k+1} | Z_k} V_{k+1}(\omega(k+1)) \right\} \\ &= E_{Z_k} \min_{\hat{x}_k} \left\{ E_{Z_{k+1} | Z_k} C(\hat{x}_k, x_k, k) + E_{Z_{k+1} | Z_k} V_{k+1}(\omega(k+1)) \right\} \\ &= E_{Z_k} V_k(\omega(k)) \end{aligned}$$

where

$$V_k(\omega(k)) = \min_{\hat{x}_k} \left\{ E_{Z_k} C(\hat{x}_k, x_k, k) + E_{Z_{k+1} | Z_k} V_{k+1}(\omega(k+1)) \right\} \quad (2.17)$$

Equations (2.15)-(2.17) hold equally for all the three types of supervision and corresponding expressions are obtained by adding superscripts on the optimal returns.

It can easily be shown that for the cost function under consideration the following inequalities hold

$$C(\hat{x}_N, x_N, N) \geq \min_{\hat{x}_N^2} C(\hat{x}_N, x_N, N) \geq \min_{\hat{x}_N} C(\hat{x}_N, x_N, N) \quad (2.18)$$

Since operations such as minimization and conditional expectation do not change the inequalities,

$$\min_{\hat{x}_N} E_{|Z_N} C(\hat{x}_N, x_N, N) \geq \min_{\hat{x}_N} E_{|Z_N} \min_{\hat{x}_N^2} C(\hat{x}_N, x_N, N) \geq \min_{\hat{x}_N} E_{|Z_N} \min_{\hat{x}_N} C(\hat{x}_N, x_N, N) \quad (2.19)$$

which can be reduced to

$$\min_{\hat{x}_N} E_{|Z_N} C(\hat{x}_N, x_N, N) \geq \min_{\hat{x}_N^1} E_{|Z_N} \min_{\hat{x}_N^2} C(\hat{x}_N, x_N, N) \geq E_{|Z_N} \min_{\hat{x}_N} C(\hat{x}_N, x_N, N) \quad (2.20)$$

Utilizing the preceding inequalities and making use of the independence of x_N^1 and x_N^2 , Eq. (2.15) yields

$$\begin{aligned} V_N^P(\omega(N)) &= \min_{\hat{x}_N} E_{|Z_N} C(\hat{x}_N, x_N, N) \\ &\geq \min_{\hat{x}_N^1} E_{|Z_N} \min_{\hat{x}_N^2} C(\hat{x}_N, x_N, N) \\ &= \min_{\hat{x}_N^1} E_{|Z_N} C(\hat{x}_N^1, x_N^1, x_N^2 = \hat{x}_N^2, N) \\ &= V_N^S(\omega(N)) \end{aligned} \quad (2.21)$$

Assume that the inequality holds at stage $k+1$, i.e.

$$V_k^P(\omega(k)) \geq V_k^S(\omega(k)) \quad \forall k \quad (2.22)$$

At stage k, from Eq. (2.17)

$$\begin{aligned}
 v_k^P(\omega(k)) &= \min_{\hat{x}_k} \left\{ E_{z_k} C(\hat{x}_k, x_k, k) + E_{z_{k+1}|z_k} v_{k+1}^P(\omega(k+1)) \right\} \\
 &\geq \min_{\hat{x}_k} \left\{ E_{z_k} C(\hat{x}_k, x_k, k) + E_{z_{k+1}|z_k} v_{k+1}^S(\omega(k+1)) \right\} \\
 &= v_k^S(\omega(k))
 \end{aligned} \tag{2.23}$$

To establish the lower bound, it is noted that Equation (2.15) yields

$$\begin{aligned}
 v_N^P(\omega(N)) &= \min_{\hat{x}_N} E_{z_N} C(\hat{x}_N, x_N, N) \\
 &= \min_{\hat{x}_N^1} E_{z_N} C(\hat{x}_N^1, x_N^1, N) + \min_{\hat{x}_N^2} E_{z_N} C(\hat{x}_N^2, x_N^2, N) \\
 &\leq \min_{\hat{x}_N^1} E_{z_N} C(\hat{x}_N^1, x_N^1, N) + \min_{\hat{x}_N^2} E_{z_N} C(\hat{x}_N^2, x_N^2, N)
 \end{aligned} \tag{2.24}$$

where $C(\hat{x}_N, x_N, N) = C(\hat{x}_N^1, x_N^1, N) + C(\hat{x}_N^2, x_N^2, N)$ because of the quadratic nature of the cost function $C(\hat{x}_N, x_N, N)$ and the independence of x_N^1 and x_N^2 .

Following the same reasoning as in Eq. (2.23) one can show

$$v_k^P(\omega(k)) < v_k^U(\omega(k)) \quad \forall k \tag{2.25}$$

In subsequent chapters, we will consider a specific partially supervised estimator, namely the probabilistic estimator and derive explicit algorithms for its implementation under various conditions.

REFERENCES

- [1] R. A. Singer, R. G. Sea, and K. B. Housewright, "Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multi-Target environments", IEEE Trans. Information Theory, Vol. IT-20, pp. 423-432, July 1974.
- [2] P. Smith and G. Buechler, "A Branching Algorithm for Discriminating and Tracking Multiple Objects", IEEE Trans. Automatic Control, Vol. AC-20, pp 101-104, Feb. 1975.
- [3] Y. Bar-Shalom and A. G. Jaffer, "Adaptive Nonlinear Filtering for Tracking with Measurements of Uncertain Origin", IEEE Conf. on Decision and Control, New Orleans, pp. 243-247, Dec. 1972.
- [4] R. A. Singer and R. G. Sea, "New Results in Optimizing Surveillance System Tracking and Data Correlation Performance in Dense Multitarget Enviroments", IEEE Trans. Automatic Control. Vol. AC-18, pp. 571-582, Dec. 1973.
- [5] R. G. Sea, "An Efficient Suboptimal Decision Procedure for Associating Sensor Data with Stored Tracked in Real-Time Surveillance Systmes", Proc. IEEE conf. Decision and Control, Miami Beach, pp. 33-37, Dec. 1971.
- [6] K. J. Astrom, "Optimal Control of Markov Processes with Incomplete State Information", J. Math. Analysis and Appl., Vol. 10, pp. 174-205, 1965.
- [7] C. C. White, "Application of Two Inequality Results for Concave Functions to a Stochastic Optimization Problem", to be published in J. Math. Analysis and Appl.

3. THE PROBABILISTIC ESTIMATOR

3.1 Introduction

As discussed in Chapter I, the optimal Bayesian estimator for tracking a target in a multi-target, multi-sensor environment is not feasible because of its exploding computational requirements. In this chapter, we will present a suboptimal scheme, namely the probabilistic estimator which uses a probabilistic judgement at each stage of the observations to determine the true hypothesis. The scheme is evaluated by comparing its performance with two other suboptimal schemes which have been presented in the literature.

As explained in Chapter II, the dynamic equation of motion of the target being tracked is assumed to be modelled by the set of Gauss-Markov difference equations

$$x_{k+1} = \phi_k x_k + \Gamma_k \omega_k$$
$$y_k = c_k x_k \quad (3.1)$$

where x_k is the $n \times 1$ state vector, and y_k is a $p \times 1$ output vector.

The m -vector ω_k is white zero mean gaussian with covariance $E\{\omega_k \omega_k^T\} = Q_k \delta_{ke}$, ϕ_k is an $n \times n$ state transition matrix and Γ_k is an $n \times m$ excitation matrix.

The observations at any stage can originate from one of three sources (hypotheses), namely, (i) noise alone (e.g., cut-off communication link), (ii) false alarm (e.g. thermal or process noise, or clutter from a vehicle not being tracked), and (iii) the vehicle being tracked. The observations are thus given by [1,2]

$$H_k^1 : z_k = u_k \quad (3.2)$$

where the $p \times 1$ vector u_k is a zero mean, white Gaussian sequence with covariance $E\{u_k u_m^T\} = R_k \delta_{km}$.

$$H_k^2 : z_k = c_k \hat{x}(k|k-1) + v_k \quad (3.3)$$

where $\hat{x}(k|k-1)$ is the one-stage prediction of x_k and v_k is a white, zero mean sequence with covariance $E\{v_k v_m\} = \Omega_k \delta_{km}$

$$H_k^3 : z_k = c_k x_k + u_k \quad (3.4)$$

The noise sequences ω_k , v_k , and u_k are mutually independent and are uncorrelated with respect to the initial state x_0 which is multivariate normally distributed of dimension n with mean μ_0 and covariance V_0 , $N_n [\mu_0, V_0]$.

It is well known that the optimal Bayes estimator for this problem involves an evergrowing memory and is hence not feasible. This is because the estimate of x_k at each stage depends on the sequence of past true hypotheses. This sequence, however, is unknown and hence the estimate must be averaged over all possible past sequences. This corresponds to the unsupervised learning problems in pattern classification. It is clear that as the number of stages increases, the total number of sequences over which the estimates must be averaged increases very rapidly. We, therefore, need to resort to suboptimal schemes which use a fixed memory. There are basically two approaches which can be taken to solve this problem. We can use a fixed number, say K , of past observations in arriving at our estimate. Alternately, we can use all past observations in our estimation scheme but choose only one (or more generally, a fixed number N) of possible past sequences over which to average.

A few such schemes have been discussed in the literature. A statistical test is used in [3,4] to eliminate the likely extraneous returns. The number of sensor returns can be further reduced by combining identical observations from the most recent N stages [1]. This is essentially an estimator based on an infinite sequence of hypotheses and a finite sequence of observations. An alternative has been studied in [5], which is essentially an estimator based on an infinite sequence of observations and a finite sequence of hypotheses. Either of them requires a tremendous amount of computation. We now present the structure of a probabilistically supervised estimator which uses a probabilistically supervised learning sequence to determine the hypothesis corresponding to each sensor return and thus eliminate extraneous returns. The estimator is compared with two other suboptimal estimators in terms of performance and computational requirements. Simulation results are presented to show the applicability of the probabilistically supervised estimator.

3.2 Probabilistically Supervised Estimator

In generating a probabilistically supervised learning sequence, one has to design a probabilistic judgement at each stage of observations, on the sensor returns [6]. This judgement is carried out in parallel on each sensor return, in such a way that each return is assigned a classification label which then serves as its true hypothesis. We can think of the scheme as consisting of a set of parallel processing units (PPU), one corresponding to each of the sensors, and a central processing unit (CPU) which combines the processed returns from the PPU's to update the estimate of the target track. On receiving the observations, each

PPU performs a random labelling test to generate a label ℓ_k corresponding to each of the three hypotheses. The label is then accepted as the true hypothesis at that stage. Depending on the label, the observation from the i^{th} sensor is modeled according to Eqs. (3.2)-(3.4) by the corresponding PPU. The algorithm for combining the observation models supplied to obtain the track estimate is discussed later.

We can now consider the generation of the random labels.

Suppose there are M sensors and that at stage k , there are n_k returns, $n_k \leq M$. Let ℓ_k^i denote the label generated by the i^{th} PPU at stage k and let $L(k)$ be the vector

$$L(k) = [\ell_k^1 \ \ell_k^2 \ \dots \ \ell_k^{n_k}]^T$$

Let L_k^M denote the set of all past labels $L(j)$, $j = 1 \dots k$.

Similarly let z_k^i denote the i^{th} sensor return at stage k and

let

$$Z(k) = [z_k^1 \ z_k^2 \ \dots \ z_k^{n_k}]^T$$

Let Z_k^M denote the set of all past observations $Z(j)$, $j=1 \dots k$.

Since there are three hypotheses, the label ℓ_k^i is a ternary random variable. When a new measurement z_{k+1}^i is received, the PPU updates the probabilities associated with the three outcomes on the basis of z_{k+1}^i , the past measurement sequence Z_k^M and the past labelling sequence L_k^M . We can then write for these probabilities

$$\begin{aligned} f(\ell_{k+1}^i) &= H_{k+1}^j | z_{k+1}^i, Z_k^M, L_k^M \\ &= \frac{f(z_{k+1}^i | \ell_{k+1}^i = H_{k+1}^j, Z_k^M, L_k^M) \cdot f(\ell_{k+1}^i = H_{k+1}^j | Z_k^M, L_k^M)}{f(z_{k+1}^i | Z_k^M, L_k^M)} \end{aligned} \quad (3.5)$$

where

$$f(z_{k+1}^i | z_k^M, L_k^M) = \sum_{j=1}^3 f(z_{k+1}^i | \ell_{k+1}^i = H_{k+1}^j, z_k^M, L_k^M) f(\ell_{k+1}^i | z_k^M, L_k^M) \quad (3.6)$$

for $i = 1, 2, \dots, n_k$; $j = 1, 2, 3$

If we assume that the hypotheses are independent from stage to stage,

we can write

$$f(\ell_{k+1}^i = H_{k+1}^j | z_k^M, L_k^M) = f(\ell_{k+1}^j = H_{k+1}^j) \quad (3.7)$$

For Markov-dependent hypotheses, we have

$$f(\ell_{k+1}^i = H_{k+1}^j | z_k^M, L_k^M) = F(\ell_{k+1}^j = H_{k+1}^j | \ell_k^i = H_k^n) \quad (3.8)$$

It only remains to determine $f(z_{k+1}^i | \ell_{k+1}^i = H_{k+1}^j, z_k^M, L_k^M)$. From the observation model, Eqs. (3.1)-(3.3), we can write

$$f(z_{k+1}^i | \ell_{k+1}^i = H_{k+1}^j, z_k^M, L_k^M) = N_p[0, R_{k+1}] \quad \text{for } j=1 \quad (3.9a)$$

$$= N_p[C_{k+1} \hat{x}_{k+1} | k (L_k^M), \Omega_{k+1} (L_k^M)] \quad \text{for } j=2 \quad (3.9b)$$

$$= N_p[C_{k+1} \hat{x}_{k+1} | k (L_k^M), C_{k+1} V_{k+1} | k (L_k^M) C_{k+1}^T + R_{k+1}] \quad \text{for } j=3 \quad (3.9c)$$

Define

$$Y_{k+1}^i \triangleq z_{k+1}^i - E\{z_{k+1}^i | L_k^M, z_k^M\} \quad (3.10)$$

which is a white sequence.

Let $Y_{k+1}^M (L_{k+1}^M)$ be a collection of the white sequences

associated with all the sensor returns.

It then follows that the equations for the sequential estimation of the state x_k are given by

$$\hat{x}_{k+1} | k (L_k^M) = \phi_k \hat{x}_k (L_k^M) \quad (3.11)$$

$$v_{k+1|k}(L_k^M) = \phi_k v_k(L_k^M) \phi_k^T + \Gamma_k Q_k \Gamma_k^T \quad (3.12)$$

$$K_{k+1}(L_{k+1}^M) = v_{k+1}(L_{k+1}^M) (C_{k+1}^M)^T (R_{k+1}^M)^{-1} \quad (3.13)$$

$$\hat{x}_{k+1}(L_{k+1}^M) = \hat{x}_{k+1|k}(L_k^M) + K_{k+1}(L_{k+1}^M) y_{k+1}(L_{k+1}^M) \quad (3.14)$$

and

$$v_{k+1}(L_{k+1}^M) = [I - K_{k+1}(L_{k+1}^M) C_{k+1}^M] v_{k+1|k}(L_k^M) \quad (3.15)$$

where

$$R_{k+1}^M = \text{diag}[R_{k+1}^1, \dots, R_{k+1}^l] \quad (3.16)$$

Eqs. (3.5)-(3.16) constitute a complete set of algorithms for updating the estimate.

3.3 Evaluation of Probabilistic Estimator

We will now briefly describe two suboptimal estimator which have been discussed in the literature for purposes for comparison with the probabilistic estimator. In the first scheme, the infinite sequence of hypotheses is approximated by a sequence of fixed size. This scheme is referred to as suboptimal estimator of finite memory. The other scheme is to eliminate the likely extraneous returns by constructing a gate test. This is referred to as a suboptimal estimator with gating.

3.3.1 Suboptimal Estimator of Finite Memory

The conditional density of one-stage predictor $f(x_{k+1} | a_{k+1}, Z_k^M)$ depends on all the past sequences a_{k+1} and requires an evergrowing amount of core memory for its implementation. The sequence $a_k = (\alpha_k, \dots, \alpha_1)$

could be approximated by a subsequence b_k taking the most recent N elements, $b_k = (\alpha_k, \dots, \alpha_{k-N+1})$. As the memory size is restricted to N , the earliest element α_{k-N+1} is not used in evaluating the conditional density of one-stage predictor. In other words, $f(x_{k+1} | a_{k+1}, Z_k^M)$ is an approximated gaussian with mean $\hat{x}_{k+1|k}(b_k^s)$ and variance $v_{k+1|k}(b_k^s)$, where $b_k^s = (\alpha_k, \dots, \alpha_{k-N+2})$.

As a consequence of the approximation, the following equations hold

$$f(x_{k+1} | b_{k+1}, Z_{k+1}^M) = N[\hat{x}_{k+1|k}(b_k^s), v_{k+1|k}(b_k^s)] \quad \text{if } \alpha_{k+1} = H_{k+1}^1 \quad (3.17a)$$

$$= N[\hat{x}_{k+1|k}(b_k^s), v_{k+1|k}(b_k^s)] \quad \text{if } \alpha_{k+1} = H_{k+1}^2 \quad (3.17b)$$

$$= N[\hat{x}_{k+1}(b_{k+1}), v_{k+1}(b_{k+1})] \quad \text{if } \alpha_{k+1} = H_{k+1}^3 \quad (3.17c)$$

$$\hat{x}_{k+1|k}(b_k^s) = \phi_k \hat{x}_k(b_k^s) \quad (3.18)$$

$$v_{k+1|k}(b_k^s) = \phi_k v_k(b_k^s) \phi_k^T + \Gamma_k Q_k \Gamma_k^T \quad (3.19)$$

$$\hat{x}_k(b_k^s) = \frac{\sum_{\alpha_{k-N+1}} \hat{x}_k(b_k) \cdot f_w(b_k | Z_k^M)}{\sum_{\alpha_{k-N+1}} f_w(b_k | Z_k^M)} \quad (3.20)$$

$$v_k(b_k^s) = \frac{\sum_{\alpha_{k-N+1}} v_k(b_k) \cdot f_w(b_k | Z_k^M)}{\sum_{\alpha_{k-N+1}} f_w(b_k | Z_k^M)} \quad (3.21)$$

$$\hat{x}_{k+1}(b_{k+1}) = \hat{x}_{k+1|k}(b_k^s) + K_{k+1}(b_{k+1})[z_{k+1}^M - C_{k+1}^M \hat{x}_{k+1|k}(b_k^s)] \quad (3.22)$$

$$K_{k+1}(b_{k+1}) = V_{k+1}(b_{k+1})(C_{k+1}^M)^T (R_{k+1}^M)^{-1} \quad (3.23)$$

$$V_{k+1}(b_{k+1}) = [I - K_{k+1}(b_{k+1})C_{k+1}^M]V_{k+1|k}(b_k^s) \quad (3.24)$$

$$\hat{x}_{k+1} = \sum_{\alpha_{k+1}} \dots \sum_{\alpha_{k-N+2}} \hat{x}_{k+1}(b_{k+1}) \cdot f_w(b_{k+1} | z_{k+1}^M) \quad (3.25)$$

$$V_{k+1} = \sum_{\alpha_{k+1}} \dots \sum_{\alpha_{k-N+2}} V_{k+1}(b_{k+1}) \cdot f_w(b_{k+1} | z_{k+1}^M) \quad (3.26)$$

The weighting coefficients are updated by

$$f_w(b_{k+1} | z_{k+1}^M) = \frac{f(z_{k+1}^M | b_{k+1}, z_k^M) f(b_{k+1} | z_k^M)}{f(z_{k+1}^M | z_k^M)} \quad (3.27)$$

where

$$f(z_{k+1}^M | b_{k+1}, z_k^M) = N_{\ell p}[0, R_{k+1}^M] \quad (3.28)$$

$$\begin{aligned} & \text{if } \alpha_{k+1} = H_{k+1}^1 \\ & = N_{\ell p}[C_{k+1}^M \hat{x}_{k+1|k}(b_k^s), \Omega_{k+1}(b_k^s)] \end{aligned} \quad (3.29)$$

$$\begin{aligned} & \text{if } \alpha_{k+1} = H_{k+1}^2 \\ & = N_{\ell p}[C_{k+1}^M \hat{x}_{k+1|k}(b_k^s), C_{k+1}^M V_{k+1|k}(b_k^s)(C_{k+1}^M)^T + R_{k+1}^M] \end{aligned} \quad (3.30)$$

where ℓ denotes the number of sensor returns at stage $k+1$, and

ℓ could vary from stage to stage.

$$f(z_{k+1}^M | z_k^M) = \sum_{\alpha_{k+1}} \dots \sum_{\alpha_{k-N+2}} f(z_{k+1}^M | b_{k+1}, z_k^M) \cdot f_w(b_{k+1} | z_k^M) \quad (3.31)$$

$$f_w(b_{k+1} | z_k^M) = f_w(\alpha_{k+1} | b_k^s, z_k^M) \cdot \sum_{\alpha_{k-N+1}} f_w(b_k | z_k^M) \quad (3.32)$$

$$f_w(\alpha_{k+1} | b_k^s, z_k^M) = P(\alpha_{k+1}) \quad \text{for independent hypotheses} \quad (3.33a)$$

$$= P(\alpha_{k+1} | \alpha_k) \quad \text{for markov-dependent hypotheses} \quad (3.33b)$$

It is noted that the suboptimal estimator of memory size N is identical to the optimal estimator under a single return and the suboptimal estimator deviates from the optimal estimator for $k > N$. The sufficient statistics at the transition stage $k = N$ are $\hat{x}_N(b_N)$ and $V_N(b_N)$, which are not valid for evaluating the density of the one-stage predictor. Because of the restriction to most recent N stages, the stage $k-N+1$ cannot be used in evaluating the density of the one-stage predictor. This can be overcome by simply averaging $\hat{x}_k(b_k)$ and $V_k(b_k)$ with respect to the weighting coefficients at stage $k-N+1$

$$\hat{x}_k(b_k^s) = \sum_{\alpha_{k-N+1}} \hat{x}_k(b_k) f_w(b_k | z_k^M)$$

and

$$V_k(b_k^s) = \sum_{\alpha_{k-N+1}} V_k(b_k) f_w(b_k | z_k^M)$$

3.3.2 Suboptimal Estimator with Gating

In this scheme, only those sensor returns which pass a certain gate test [3,5] are considered for updating a particular track. Note that the innovations [7] corresponding to the correct return at stage k is approximately normally distributed with zero mean and variance

$C_k V_{k|k-1} C_k^T + R_k$. It is well known that

$$\| z_k^1 - C_k \hat{x}_{k|k-1} \|^2_{(C_k V_{k|k-1} C_k^T + R_k)^{-1}} \quad (3.34)$$

has a p degrees of freedom chi-square distribution. The gate test is to select the threshold in such a way that the sensor return z_k^1 for which the quantity in Eq. (3.34) falls in the right tail is rejected in order to eliminate the extraneous returns. Hence the number of validated returns could be limited and utilized in track updating.

It is assumed that there is a nonzero probability of incorrect correlation. Let the set of validated returns at stage $k+1$ be

$$z_{k+1}^M = \{z_{k+1}^1, \dots, z_{k+1}^\ell\} \quad (3.35)$$

and the accumulated validated returns up to stage $k+1$ be

$$z_{k+1}^M = \{z_{k+1}^M, \dots, z_1^M\} \quad (3.36)$$

Define at each stage, the events

$$\Lambda_{k+1}^0 = \{\text{proper return is detected but it is not correctly correlated}\} \cup \{\text{proper return is not detected}\}$$

and

$$\Lambda_{k+1}^i = \{z_{k+1}^i \text{ is correctly correlated at stage } k+1\}$$

where $i = 1, 2, \dots, \ell$; and $k = 1, 2, \dots$. Assume that Λ_{k+1}^i are mutually exclusive and equally likely events, and that $\{z_{k+1}^1, \dots, z_{k+1}^\ell\}$ are jointly independent and identically distributed. The probability of event Λ_{k+1}^0 is, for simplicity, assumed constant and known a priori.

The minimum variance estimate is

$$\begin{aligned} \hat{x}_{k+1} &= \int_{\Omega} x_{k+1} f(x_{k+1} | z_{k+1}^M) dx_{k+1} \\ &= \sum_{i=0}^{\ell} \int_{\Omega} x_{k+1} f(x_{k+1} | z_{k+1}^M, \Lambda_{k+1}^i) dx_{k+1} \cdot f_w(\Lambda_{k+1}^i | z_{k+1}^M) \\ &= \sum_{i=0}^{\ell} \hat{x}_{k+1}(\Lambda_{k+1}^i) \cdot f_w(\Lambda_{k+1}^i | z_{k+1}^M) \end{aligned} \quad (3.37)$$

The variance associated with the above estimate is obtained as follows

$$\begin{aligned}
 V_{k+1} &= \int_{\Omega} (x_{k+1} - \hat{x}_{k+1})(x_{k+1} - \hat{x}_{k+1})^T f(x_{k+1} | Z_{k+1}^M) dx_{k+1} \\
 &= \sum_{i=0}^{\ell} V_{k+1}(\Lambda_{k+1}^i) f_w(\Lambda_{k+1}^i | Z_{k+1}^M) \\
 &\quad + \sum_{i=0}^{\ell} \hat{x}_{k+1}(\Lambda_{k+1}^i) \hat{x}_{k+1}^T(\Lambda_{k+1}^i) \cdot f_w(\Lambda_{k+1}^i | Z_{k+1}^M) - \hat{x}_{k+1} \hat{x}_{k+1}^T
 \end{aligned} \tag{3.38}$$

It is to be noted that the variance is the sum of weighted variance plus a positive semidefinite matrix to incorporate the effect of the extraneous measurements [3].

The conditional a posteriori density is calculated as

$$f(x_{k+1} | \Lambda_{k+1}^i, Z_{k+1}^M) = \frac{f(z_{k+1}^M | x_{k+1}, \Lambda_{k+1}^i, Z_k^M) \cdot f(x_{k+1} | \Lambda_{k+1}^i, Z_k^M)}{f(z_{k+1}^M | \Lambda_{k+1}^i, Z_k^M)} \tag{3.39}$$

The above expression can be simplified by making use of the assumption on Λ_{k+1}^i , as

$$f(x_{k+1} | \Lambda_{k+1}^i, Z_{k+1}^M) = \frac{f(z_{k+1}^i | x_{k+1}, \Lambda_{k+1}^i, Z_k^M) f(x_{k+1} | \Lambda_{k+1}^i, Z_k^M)}{f(z_{k+1}^i | \Lambda_{k+1}^i, Z_k^M)} \tag{3.40}$$

In evaluating Eq. (3.40), the following should be noted:

1. The conditional density of one-stage predictor is independent of Λ_{k+1}^i , and is equal to $N[\hat{x}_{k+1} | k, V_{k+1} | k]$.
2. $f(z_{k+1}^i | x_{k+1}, \Lambda_{k+1}^i, Z_k^M) = N_p[C_{k+1} x_{k+1}, R_{k+1}]$ for $i = 1, 2, \dots, \ell$; and either $N_p[0, R_{k+1}]$ or $N_p[C_{k+1} \hat{x}_{k+1} | k, \Omega_{k+1}]$ for $i=0$.

It is clear from the above that the conditional a posteriori density can therefore be written as:

$$f(x_{k+1} | \Lambda_{k+1}^i, Z_{k+1}^M) = N_n[\hat{x}_{k+1}(\Lambda_{k+1}^i), v_{k+1}(\Lambda_{k+1}^i)]$$

for $i = 1, 2, \dots, \ell$ (3.41)

$$= N_n[\hat{x}_{k+1|k}, v_{k+1|k}]$$

for $i=0$ (3.42)

together with the following equations for updating the sufficient statistics,

$$\hat{x}_{k+1}(\Lambda_{k+1}^i) = \hat{x}_{k+1|k} + K_{k+1}(\Lambda_{k+1}^i)[z_{k+1}^i - C_{k+1}\hat{x}_{k+1|k}]$$

(3.43)

$$K_{k+1}(\Lambda_{k+1}^i) = v_{k+1|k} C_{k+1}^T [C_{k+1} v_{k+1|k} C_{k+1}^T + R_{k+1}]^{-1}$$

(3.44)

$$v_{k+1}(\Lambda_{k+1}^i) = [I - K_{k+1}(\Lambda_{k+1}^i) C_{k+1}] v_{k+1|k}$$

(3.45)

As expected the gain $K_{k+1}(\Lambda_{k+1}^i)$ is independent of Λ_{k+1}^i and is the same for each validated return.

The weighting coefficients are calculated as follows:

Since

$$f_w(\Lambda_{k+1}^i | Z_{k+1}^M) = \frac{f(z_{k+1}^M, \Lambda_{k+1}^i | Z_k^M)}{f(z_{k+1}^M | Z_k^M)}$$

$$= \frac{f(z_{k+1}^i, \Lambda_{k+1}^i | Z_k^M)}{f(z_{k+1}^i | Z_k^M)}$$

$$= f(\Lambda_{k+1}^i | z_{k+1}^i, Z_k^M), i=1, 2, \dots, \ell$$

(3.46)

It follows from the preceding that

$$f_w(\Lambda_{k+1}^i | z_{k+1}^i, Z_k^M) = f_w(\Lambda_{k+1}^j | z_{k+1}^j, Z_k^M)$$

$$= \frac{1 - P(\Lambda_{k+1}^0 | Z_{k+1}^M)}{\ell}$$

(3.47)

where $P(\Lambda_{k+1}^0 | Z_{k+1}^M)$ is known a priori [3].

Equations (3.37), (3.38), and (3.41)-(3.47) constitute the necessary set of equations for recursively updating the sufficient statistics.

3.4 Simulations

In order to evaluate the performance of the probabilistically supervised estimator and to compare it with the other estimators, the same tracking example given in [1] was simulated. The parameters of the aircraft being tracked are given as follows

$$x_k = \begin{bmatrix} x - \text{position at stage } k \\ x - \text{speed at stage } k \\ y - \text{position at stage } k \\ y - \text{speed at stage } k \end{bmatrix} \quad (4 \times 1)$$

$$\omega_k = \begin{bmatrix} \text{change in } x\text{-speed between stage } k \text{ and stage } k+1 \\ \text{change in } y\text{-speed between stage } k \text{ and stage } k+1 \end{bmatrix} \quad (2 \times 1)$$

$$\phi_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4 \times 4)$$

$$\Gamma_k = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (4 \times 2)$$

$$\text{and } Q_k = \begin{bmatrix} q_{11}(k) & 0 \\ 0 & q_{22}(k) \end{bmatrix} \quad (2 \times 2)$$

where T is the sampling interval

$$C_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2 \times 4)$$

$$R_k = \begin{bmatrix} \gamma_{11}(k) & 0 \\ 0 & \gamma_{22}(k) \end{bmatrix} \quad (2 \times 2)$$

For simplicity, the following values are used

$$T = 2$$

$$\rho = 0$$

$$n_{TF} = 0.01$$

$$x_0, \text{ multivariate normally distributed with dimension four, } N_4[0, 0.01].$$

$$Q_k = \text{diag}[0, 0]$$

and

$$R_k = \text{diag} [0.00828, 0.00828]$$

Simulation results are presented in the form of tables for the mean square error associated with the x-position. Similar results were also obtained for the other states.

The simulations consist of a comparison between the three estimators presented in this chapter, namely Suboptimal Estimator with Finite Memory (SFE), Suboptimal Estimator with Grating (SGE) and the Probabilistically Supervised Estimator (PSE). One hundred simulations under a single return were averaged to obtain the mean square errors. The probability of the hypotheses were $P(H_k^3) = 1$, $P(H_k^1) = P(H_k^2) = 0$. The results are shown in Table 3.1 which also shows the mean square error for the optimal Bayesian Estimator (OBE). For the simulations, an infinite gate size, unity detection probability of sensor and $P(\Lambda_{k+1}^0 | Z_{k+1}^M) = 0$ were assigned to SGE and a memory length of two stages to SFE. The simulations were carried to only five stages because of the computational requirements associated with OBE. The performance of all the estimators is equally acceptable under these conditions.

Table 3.2 shows the results for the case $P(H_k^1) = P(H_k^2) = 0.1$ and $P(H_k^3)=0.8$. The decrease in MSE for the PSE is not as fast as with the OBE and SFE. However, as shown in Table 3.3 the PSE has decided advantages in terms of computation requirements.

The second part of the simulations consist of a comparison between PSE and SFE. A one-hundred-step process under five ground sensors was simulated on the same model. Ten simulations were averaged to obtain the mean square errors and variances. In Tables 3.4, 3.5 and 3.6, it is shown that the performance of PSE is better than that of SFE in all the possible cases. This fact becomes more pronounced in the most uncertain environments in which $P(H_k^1)=0.2$, $P(H_k^2)=P(H_k^3)=0.4$.

For the case of dependent hypotheses transition probabilities $P(H_{k+1}^j | H_k^i)$ were chosen as

	$j = 1$	2	3
$i = 1$	0.8	0.1	0.1
2	0.1	0.8	0.1
3	0.1	0.1	0.8

The process was started initially with equally likely probabilities of hypotheses. The results shown in Table 3.7 indicate that the performance of PSE is quite acceptable in this case.

TABLE 3.1

MSE of PSE, OBE, SFE, and SGE; $P(H_k^3) = 1$

Stage	PSE	OBE	SFE	SGE
1	.02539	.03300	.03300	.02675
2	.01790	.02109	.02109	.02331
3	.03262	.03301	.03301	.03466
4	.03333	.03202	.03202	.03768
5	.03079	.03107	.03107	.03195

TABLE 3.2

MSE of PSE, OBE, SFE, and SGE; $P(H_k^1)=P(H_k^2)=0.1$ and $P(H_k^3)=0.8$

Stage	PSE	OBE	SFE	SGE
1	.34901	.20144	.20144	.17598
2	.13036	.12237	.12237	.22925
3	.18304	.11934	.14284	.26749
4	.21962	.08673	.11383	.28088
5	.14741	.07520	.09133	.47075

TABLE 3.3

Percentage of Time and Core Locations

	Time ¹	Core 2 Locations
PSE	6	48
OBE	100	100
SFE	16	65
SGE	6	59

1. Percentage of the computer time required by the optimal filter.
2. Percentage of the core locations required by the optimal filter.

TABLE 3.4

MSE of PSE and SFE; $p(H_k^3) = 1$

STAGE	PSE	SFE
10	.00153	.01967
20	.00101	.00589
30	.00022	.00294
40	.00025	.00196
50	.00019	.00136
60	.00017	.00101
70	.00016	.00077
80	.00019	.00058
90	.00014	.00042
100	.00010	.00033

-TABLE 3.5

MSE of PSE and SFE; $p(H_k^1) = p(H_k^2) = 0.1$ and $p(H_k^3) = 0.8$

STAGE	PSE	SFE
10	.02071	.01439
20	.00283	.02314
30	.00159	.01425
40	.00065	.02270
50	.00037	.03167
60	.00025	.05524
70	.00021	.07709
80	.00022	.07149
90	.00019	.09386
100	.00012	.10536

TABLE 3.6

MSE of PSE and SFE; $p(H_k^1) = 0.2$ and $p(H_k^2) = p(H_k^3) = 0.4$

STAGE	PSE	SFE
10	.69687	.10443
20	1.13143	.09342
30	.45878	.08765
40	.37909	.19294
50	.32131	.26892
60	.22679	.48263
70	.19707	.59716
80	.16257	.75582
90	.11455	.87802
100	.10383	.98750

TABLE 3.7

MSE of PSE and SFE; Markov-dependent hypotheses

STAGE	PSE	SFE
10	.13949	1.44858
20	.09721	1.37159
30	.10936	1.06421
40	.15072	.76816
50	.23789	.74486
60	.25945	.89461
70	.28760	1.14302
80	.30237	1.09317
90	.34798	1.30457
100	.39394	1.97090

REFERENCES

- [1] R.A. Singer, R.G. Sea, and K.B. Housewright, "Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multitarget Environments", IEEE Trans. Information Theory, Vol. IT-20, pp. 423-432, July 1974.
- [2] R.A. Singer and R.G. Sea, "New Results in Optimizing Surveillance System Tracking and Data Correlation Performance in Dense Multitarget Environments", IEEE Trans. Automatic Control, Vol. AC-18, pp. 571-582, Dec. 1973.
- [3] Y. Bar-Shalom and A.G. Jaffer, "Adaptive Nonlinear Filtering for Tracking with Measurements of Uncertain Origin", IEEE Conf. on Decision and Control, New Orleans, pp. 243-247, Dec. 1972.
- [4] R.G. Sea, "An Efficient Suboptimal Decision Procedure for Associating Sensor Data with Stored Tracks in Real-Time Surveillance Systems", Proc. IEEE Conf. Decision and Control, Miami Beach, pp. 33-37, Dec. 1971.
- [5] A.G. Jaffer and S.C. Gupta, "On Estimation of Discrete Process Under Multiplicative and Additive Noise Conditions", Information Science, Vol. 3, pp. 267-276, 1971.
- [6] A.K. Agrawala, "Learning with a Probabilistic Teacher", IEEE Trans. Information Theory, Vol. IT-16, pp. 373-379, July 1970.
- [7] T. Kailath, "An Innovations Approach to Least Squares Estimation - Part I: Linear Filtering in Additive White Noise", IEEE Trans. Automatic Control, Vol. AC-13, pp. 646-655, Dec. 1968.

4. COUPLED RISK FORMULATION

4.1 Introduction

It has been pointed out earlier, that in all problems involving uncertainty of signal reception, the operations of detection and estimation are intimately coupled. The optimal processing of sensor returns requires the mutual coupling of the costs associated with both operations. Signal estimation is strongly influenced by the detector's outcome. A Bayes' optimum theory of joint detection and estimation of signals in noise has been studied in [1-4] by using a generalized cost function which reflects the mutual coupling between the joint operations.

An estimator based on minimizing such a coupled risk may perform better than the estimator in which the two operations are treated independently. This improvement in performance, if any, will however be at the cost of added complexity in the receiver. In this chapter a simultaneous detection and estimation scheme is developed for situations where the signal evolves as a stochastic process and uncertainty characterizing the presence of the process in the observation varies in the fashion discussed in Chapter 2. Since the optimal processor again requires an ever-growing memory, a finite memory suboptimal scheme which minimizes the coupled risk is developed. The performance of this suboptimal processor is compared with that of the probabilistic estimator introduced in the previous chapter.

4.2 Coupled Risk Formulation

There are two different kinds of risks associated with joint detection and estimation: uncoupled risk and coupled risk. In the uncoupled risk case the minimization of compound risk is equivalent to the minimization of the individual risks associated with the detection and estimation problems respectively. The minimization of the decision risk leads to the

familiar likelihood ratio test [1]. The minimization of estimation risk, however, yields the problem in which the estimator's performance is to be optimized under uncertainty [5,6]. In coupled risk case the joint operations of detection and estimation should be coupled and the costs associated with these two operations chosen to reflect the interaction between them. Hence, the total risk of joint detection and estimation can not be expressed as the sum of two independent components. A simultaneous minimization of the total risk with respect to the decision and the estimation rules is necessary. Joint detection and estimation of coupled risk has been studied in [13,15] under a single return. This section is devoted to the formulation of the problem of optimizing a coupled risk function under multi-sensor returns.

To reflect the coupling between detection and estimation, let $C_{ij}(\hat{x}_k, x_k)$ be defined as the cost of choosing the hypothesis η_k^i and making an estimate \hat{x}_k when, in fact, the hypothesis η_k^j is true and x_k is the true value of the signal. In other words, $C_{ij}(\hat{x}_k, x_k)$, the cost at stage k , can be expressed as

$$C_{ij}(\hat{x}_k, x_k) = L_{ij} + F_{ij} \cdot ||x_k - \hat{x}_k||^2 \quad (4.1)$$

where L_{ij} and F_{ij} are chosen appropriately, to indicate the interaction between the joint operations of detection and estimation.

The corresponding coupled risk is given by

$$R = \sum_{j=1}^3 \sum_{i=1}^3 \int_{\Delta_i} \int_{\Omega} C_{ij}(\hat{x}_k, x_k) f(x_k, H_k^j, Z_k^M) dx_k dZ_k^M \quad (4.2)$$

Here Δ is the observation space and Δ_i is the region in which the decision η_k^i is true so that $\Delta = \sum_{i=1}^3 \Delta_i$, Ω is the signal space, and Z_k^M refers to all the past sensor returns accumulated up to stage k .

The optimal decision is then to decide in favor of hypothesis η_k^j when the following inequality

$$\sum_{i=1}^3 \int_{\Omega} C_{ji}(\hat{x}_k, x_k) f(x_k, H_k^1, Z_k^M) dx_k < \sum_{i=1}^3 \int_{\Omega} C_{li}(\hat{x}_k, x_k) \cdot f(x_k, H_k^1, Z_k^M) dx_k \quad (4.3)$$

is satisfied for all $l \neq j$; $j, l = 1, 2, 3$.

The estimator that minimizes the coupled risk of Eq. (4.2) is defined as the optimal estimator under multi-sensor, uncertain environments. The corresponding optimal detector is given by Eq. (4.3). Again it is not feasible to implement the optimal estimator and the optimal detector, because of the need for evergrowing memories. Therefore the suboptimal alternative must be resorted to.

4.3 Suboptimal Estimator and Detector

In this section, a suboptimal estimator with finite memory requirements for the cost function of Eq. (4.1) is developed. It is understood that if the decision is that either η_k^1 or η_k^2 is true, there is no updating in the estimate. Hence, there is no cost incurred for estimation when either η_k^1 or η_k^2 is decided. In Eq. (4.1) x_k appears only when, in fact, η_k^3 is true ($j=3$).

The cost associated with the joint operations is described by the following equations,

$$\begin{aligned} C_{11}(\hat{x}_k, x_k) &= L_{11}, \\ C_{12}(\hat{x}_k, x_k) &= L_{12}, \\ C_{13}(\hat{x}_k, x_k) &= L_{13} + F_{13}^T x_k x_k^T, \\ C_{21}(\hat{x}_k, x_k) &= L_{21}, \\ C_{22}(\hat{x}_k, x_k) &= L_{22}, \\ C_{23}(\hat{x}_k, x_k) &= L_{23} + F_{23}^T x_k x_k^T, \\ C_{31}(\hat{x}_k, x_k) &= L_{31} + F_{31}^T \hat{x}_k \hat{x}_k^T, \\ C_{32}(\hat{x}_k, x_k) &= L_{32} + F_{32}^T \hat{x}_k \hat{x}_k^T, \end{aligned}$$

and
$$C_{33}(\hat{x}_k, x_k) = L_{33} + F_{33} (x_k - \hat{x}_k)^T (x_k - \hat{x}_k)$$

The estimate \hat{x}_k is obtained by minimizing Eq. (4.2) with respect to \hat{x}_k . This is equivalent to minimizing

$$\sum_{j=1}^3 \int_{\Delta_3} \int_{\Omega} C_{3j}(\hat{x}_k, x_k) f(x_k, H_k^j, Z_k^M) dx_k dZ_k^M \quad (4.4)$$

with respect to \hat{x}_k . Since the cost $C_{ij}(\hat{x}_k, x_k)$ is of quadratic form, the minimization of Eq. (4) yields the weighted conditional mean estimate

$$\hat{x}_k^{**} = \frac{F_{33} \int_{\Omega} x_k \cdot f(x_k | \alpha_k = H_k^3, Z_k^M) dx_k \cdot f_w(\alpha_k = H_k^3 | Z_k^M)}{\sum_{i=1}^3 F_{3i} f_w(\alpha_k = H_k^i | Z_k^M)} \quad (4.5)$$

$$\frac{\sum_{\alpha_{k-1}} \dots \sum_{\alpha_1} F_{33} \hat{x}_k^3(a_k) f_w^3(a_k | Z_k^M)}{\sum_{i=1}^3 \sum_{\alpha_{k-1}} \dots \sum_{\alpha_1} F_{3i} f_w^i(a_k | Z_k^M)} \quad (4.6)$$

where

$$\hat{x}_k^3(a_k) = \hat{x}_k(\alpha_k = H_k^3, \alpha_{k-1}, \dots, \alpha_1). \quad (4.7)$$

$$f_w^1(a_k | Z_k^M) = f_w(\alpha_k = H_k^1, \alpha_{k-1}, \dots, \alpha_1 | Z_k^M), \quad (4.8)$$

and

$$a_k = (\alpha_k, \alpha_{k-1}, \dots, \alpha_1).$$

The corresponding suboptimal estimator of finite memory which uses all the observations from the most recent N stages is given as follows

$$\hat{x}_k^* = \frac{\sum_{\alpha_{k-1}} \dots \sum_{\alpha_{k-N+1}} F_{33} \hat{x}_k^3(b_k) f_w^3(b_k | Z_k^M)}{\sum_{i=1}^3 \sum_{\alpha_{k-1}} \dots \sum_{\alpha_{k-N+1}} F_{3i} f_w^i(b_k | Z_k^M)} \quad (4.9)$$

where $\hat{x}_k^3(b_k) = \hat{x}_k(\alpha_k = H_k^3, \alpha_{k-1}, \dots, \alpha_{k-N+1}),$

$$f_w^1(b_k | Z_k^M) = f_w(\alpha_k = H_k^1, \alpha_{k-1}, \dots, \alpha_{k-N+1} | Z_k^M),$$

and $b_k = (\alpha_k, \alpha_{k-1}, \dots, \alpha_{k-N+1})$.

It is to be noted that the set of necessary equations for updating the estimate, the variance, and the weighting coefficients is identical to those derived in Chapter 3.

The decision rule is postulated in Eq. (4.3), and the observation space Δ is partitioned into three disjoint regions.

Region Δ_1 in which η_k^1 is true is true is bounded by the following inequalities.

$$\begin{aligned} & \sum_{i=1}^3 \int_{\Omega} C_{1i}(\hat{x}_k, x_k) f(x_k, H_k^i, Z_k^M) dx_k \\ & < \sum_{i=1}^3 \int_{\Omega} C_{\ell i}(\hat{x}_k, x_k) f(x_k, H_k^i, Z_k^M) dx_k, \quad \ell = 2, 3 \end{aligned} \quad (4.10)$$

In order to evaluate these inequalities one has to substitute the cost defined in the beginning of this section into Eq. (4.10) and recognize the definition of the covariance matrix

$$\begin{aligned} & \int_{\Omega} x_k x_k^T f(x_k | \alpha_k = H_k^i, \alpha_{k-1}, \dots, \alpha_{k-N+1}, Z_k^M) dx_k \\ & = V_k^i(b_k) + \hat{x}_k^i(b_k) \cdot \hat{x}_k^i(b_k)^T \end{aligned} \quad (4.11)$$

Thus

$$\begin{aligned} & \int_{\Omega} x_k x_k^T f(x_k | \alpha_k = H_k^i, \alpha_{k-1}, \dots, \alpha_{k-N+1}, Z_k^M) dx_k \\ & = T_r V_k^i(b_k) + \hat{x}_k^i(b_k)^T \cdot \hat{x}_k^i(b_k) \end{aligned} \quad (4.12)$$

where $V_k^i(b_k) = V_k(\alpha_k = H_k^i, \alpha_{k-1}, \dots, \alpha_{k-N+1})$, and $T_r V_k^i(b_k)$ denotes the trace of the matrix $V_k(b_k)$.

Define

$$A = \sum_{\alpha_{k-1}}^{\Delta} \dots \sum_{\alpha_{k-N+1}} \left\{ (L_{11} - L_{21}) \cdot f_w^1(b_k | Z_k^M) + (L_{12} - L_{22}) f_w^2(b_k | Z_k^M) \right\}$$

$$+ (L_{13}-L_{23})f_w^3(b_k|z_k^M) + (F_{13}-F_{23})(T_r v_k^3(b_k) + \hat{x}_k^3(b_k)^T \cdot \hat{x}_k^3(b_k)) \cdot f_w^3(b_k|z_k^M)\}, \quad (4.13)$$

$$B \triangleq \sum_{\alpha_{k-1}} \dots \sum_{\alpha_{k-N+1}} \left\{ (L_{11}-L_{31})f_w^1(b_k|z_k^M) + (L_{12}-L_{32})f_w^2(b_k|z_k^M) + (L_{13}-L_{33})f_w^3(b_k|z_k^M) + (F_{13}-F_{33})(T_r v_k^3(b_k) + \hat{x}_k^3(b_k)^T \hat{x}_k^3(b_k)) \cdot f_w^3(b_k|z_k^M) - (F_{13}f_w^1(b_k|z_k^M) + F_{32}f_w^2(b_k|z_k^M) + F_{33}f_w^3(b_k|z_k^M)) \cdot (\hat{x}_k^*)^T \cdot (\hat{x}_k^*) + F_{33} \cdot (\hat{x}_k^3(b_k)^T \hat{x}_k^* + (\hat{x}_k^*)^T \hat{x}_k^3(b_k)) \cdot f_w^3(b_k|z_k^M) \right\}, \quad (4.14)$$

and

$$C \triangleq B - A. \quad (4.15)$$

It is noted, from the definition of cost incurred for joint detection and estimation, that the estimate \hat{x}_k in $C_{ij}(\hat{x}_k, x_k)$ should be essentially replaced by \hat{x}_k^* of Eq. (4.9).

Upon substituting Eqs. (4.13) and (4.14) into Eq. (4.3), one can determine the boundary of region Δ_1 . It can easily be shown that the region Δ_1 is bounded by $A < 0$ and $B < 0$. It similarly follows from Eq. (4.3) that the region Δ_2 is bounded by $A > 0$ and $C < 0$, and the region Δ_3 is bounded by $B > 0$ and $C > 0$.

It is noted that the decision boundaries determined by Eqs. (4.13) - (4.15) are equivalent to the boundaries determined by setting up the modified generalized likelihood ratios [15].

The procedure to generate the optimal decision is summarized in the following:

Step 1. Evaluate A and B, respectively using Eqs. (4.13) and (4.14)

Step 2. Validate one of the following three statements:

(i) $A < 0$ and $B < 0$, decide η_k^1

(ii) $A > B$ and $A > 0$, decide η_k^2

(iii) $B > A$ and $B > 0$, decide η_k^3

Step 3. Calculate the optimal estimate

(i) Either η_k^1 or η_k^2 is true, estimate is not updated;

$$\hat{x}_k = \hat{x}_{k|k-1}$$

(ii) η_k^3 is true, estimate is updated; $\hat{x}_k = \hat{x}_k^*$.

This scheme will now be compared to the PSE scheme introduced in Chapter 3. The coupled risk formulation is not particularly meaningful to the PSE, since the estimate is updated no matter what hypothesis is true. Thus, rather than comparing the two schemes in terms of the coupled cost, it would appear more appropriate to compare the two in terms of the percentage of wrong detection and the variance of the estimator.

4.4 Simulation Results

The message model and the observation model chosen for joint detection and estimation are the same as those of Chapter 3. The values of matrices L and F in Eq. (4.1) which are assigned to the cost of making decision and the cost of making estimation respectively, are

$$L = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad F = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

for the suboptimal estimator and detector of finite memory. The values of matrices L and F also indicate the interaction of coupling of the joint operations, which is weighted equally in this case.

A one-hundred-step process was simulated and ten runs were averaged to obtain the results such as variance and percentage of wrong detection under five ground sensors. It is assumed that the a priori probabilities in the case of independent hypotheses and the transition matrix in the case of markov-dependent hypotheses are known to the receiver.

TABLE 4.1

Variances of PSE and SFE; $P(H_k^3) = 1.$

STAGE	PSE	SFE
10	.00056	.00056
20	.00030	.00030
30	.00021	.00021
40	.00016	.00016
50	.00013	.00013
60	.00011	.00011
70	.00009	.00009
80	.00008	.00008
90	.00007	.00007
100	.00007	.00007

TABLE 4.2a

Variances of PSE and SFE; $P(H_k^1)=P(H_k^2)=.1$ and $P(H_k^3)=.8$

STAGE	PSE	SFE
10	.00083	.00085
20	.00041	.00043
30	.00027	.00029
40	.00021	.00021
50	.00016	.00017
60	.00013	.00014
70	.00012	.00012
80	.00010	.00011
90	.00009	.00009
100	.00008	.00008

TABLE 4.2b

% Wrong Detection of PSE and SFE; $P(H_k^1)=P(H_k^2)=.1$ and $P(H_k^3)=.8$

STAGE	PSE	SFE
10	22	20
20	26	20
30	14	18
40	24	16
50	22	14
60	10	16
70	42	14
80	10	24
90	16	14
100	8	32

TABLE 4.3a

Variances of PSE and SFE; $P(H_k^1)=.2$ and $P(H_k^2)=P(H_k^3)=.4$

STAGE	PSE	SFE
10	.00239	.19106
20	.00109	.03719
30	.00065	.00131
40	.00049	.00071
50	.00035	.00054
60	.00029	.00043
70	.00026	.00036
80	.00023	.00030
90	.00021	.00026
100	.00019	.00022

TABLE 4.3b

% Wrong Detection of PSE and SFE; $P(H_k^1) = .2$ and $P(H_k^2) = P(H_k^3) = .4$

STAGE	PSE	SFE
10	46	60
20	50	64
30	46	56
40	48	50
50	38	58
60	34	56
70	36	56
80	32	64
90	40	50
100	40	54

TABLE 4.4a

Variances of PSE and SFE; Markov-dependent Hypotheses

STAGE	PSE	SFE
10	.00385	3.48371
20	.00090	5.25671
30	.00102	5.17088
40	.00058	4.11383
50	.00055	3.63712
60	.00041	2.86978
70	.00030	2.31418
80	.00025	1.46436
90	.00023	1.02339
100	.00021	.70642

TABLE 4.4b

% Wrong Detection of PSE and SFE, Markov-dependent Hypotheses

STAGE	PSE	SFE
10	48	64
20	36	74
30	36	74
40	26	66
50	30	62
60	38	58
70	32	62
80	38	56
90	42	64
100	48	66

TABLE 4.5

Overall % Wrong Detection of PSE and SFE

CASE	PSE	SFE
$P(H_k^3) = 1$	0.	0.
$P(H_k^1) = P(H_k^2) = .1, P(H_k^3) = .8$	18.16	21.36
$P(H_k^1) = .2, P(H_k^2) = P(H_k^3) = .4$	42.30	57.90
Markov-dependent	35.18	63.48

The variances shown in the figures of this chapter are the variances associated with the first component of the state vector.

For the case of $P(H_k^3)=1$, both schemes perform equally well and have zero percentage of wrong detection. The averaged variances at particular stages are shown in Tables 4.1, 4.2a, and 4.3a for the case of independent hypotheses. As expected the variance of PSE is lower than that of SFE. The improvement of performance of PSE over that of SFE becomes pronounced in the markov-dependent hypotheses case shown in Table 4.4a. The averaged percentage of wrong detection at particular stages are shown in Tables 4.2b, and 4.3b for the case of independent hypotheses and Table 4.4b for markov-dependent hypotheses. The overall averages of percentage of wrong detection are shown in Table 4.5. It is expected that the percentage of wrong detection for PSE is lower than that of SFE, and it is approximately fifty percent lower than SFE for the case in which the hypotheses switch in markov manner.

The structure of SFE developed in this chapter is very complicated in the sense that it takes a larger amount of computational requirements such as execution time and core locations as compared to those taken by the PSE. The simulation results clearly show the superiority of the PSE as compared to the SFE designed using the coupled risk formulation.

REFERENCES

- [1] Middleton, D. and Esposito, R., "Simultaneous Optimum Detection and Estimation of Signals in Noise, "IEEE Trans. Information Theory, Vol. IT-14, pp. 434-444, May 1968.
- [2] Esposito, R., "On a Relation between Detection and Estimation in Decision Theory," Information and Control, Vol. 12, pp. 116-120, Feb. 1968.
- [3] Jaffer, A.G. and Gupta, S.C., "Coupled Detection-Estimation of Gaussian Processes in Gaussian Noise," IEEE Trans. Information Theory, Vol. IT-18, pp. 106-110, Jan. 1972.
- [4] Birdsall, T.G. and Gobien, J.O., "Sufficient Statistics and Reproducing Densities in Simultaneous Sequential Detection and Estimation," IEEE Trans. Information Theory, Vol. IT-19, pp. 760-768, Nov. 1973.
- [5] Nahi, N.E., "Optimal Recursive Estimation with Uncertain Observation," IEEE Trans. Information Theory, Vol. IT-15, pp. 457-462, July 1969.
- [6] Jaffer, A.G. and Gupta, S.C., "Recursive Bayesian Estimation with Uncertain Observation," IEEE Trans. Information Theory, Vol. IT-17, pp. 614-616, Sept. 1971.

5. TRACKING OVER FADING CHANNELS

A problem encountered in many applications is that of distortion of the transmitted signal by the communication channel. The distortion introduced by the channel can be characterized by multiplicative disturbances which are responsible for such phenomena as fading, dispersion and multipath [1]. As can be expected, the presence of such phenomena complicates the problem of target tracking in a multisensor environment. Explicit algorithms for the probabilistic estimator for this problem have been derived and its performance evaluated by simulations which confirm the advantages probabilistic estimator. Details of the algorithms and the simulation results are given in [2] which has been included as Appendix B of this report.

REFERENCES

- [1] H. L. Van Trees, "Detection, Estimation and Modulation Theory", John Wiley and Sons, Inc., New York, 1971.
- [2] F. S. Chang and M. D. Srinath, "Target Tracking Over Fading Channels," IEEE Trans. Communications, Vol. COM-24, No. 4, April 1976, pp. 432-438.

6. ESTIMATION USING QUANTIZED MEASUREMENTS

6.1 Introduction

In many practical situations, the observations are presented to the receiver in quantized form. In general, measurements will be quantized if they originate from digital sensors, in digital transmission systems or when processed by digital computers. In this chapter we will derive algorithms for the problem of tracking a target in a multi-target, multi-sensor environment, when the observations are in quantized form, using the probabilistic estimator approach.

It is well known [1] that the problem of estimation with quantized measurements can be considered to be equivalent to the estimation of signals in the presence of additive white noise.

For a measurement equation of the form

$$z = h(x, v) \quad (6.1)$$

the conditional mean (minimum mean square estimate) of any function $f(x)$ given quantized measurements $z \in A$ where A is some region (e.g. a hypercube) can be written as

$$E\{f(x) | z \in A\} = E\{E\{f(x) | z\} | z \in A\} \quad (6.2)$$

Thus the required conditional expectation is obtained in two steps: (1) find $E(f(x) | z)$: this is the usual goal of estimation with unquantized measurements; (2) find the expectation of $E(f(x) | z)$ conditioned on $z \in A$.

For the case when

$$\begin{aligned} x &= N(\bar{x}, M) \\ v &= N(0, R) \\ z &= Hx + v \end{aligned} \quad E(xv^T) = 0 \quad (6.3)$$

where $N(\cdot, \cdot)$ denotes a normal distribution whose mean and covariance are the first and second arguments, respectively, x is an n -component vector, z and v are m -component vectors, and H is an $m \times n$ matrix, the following expressions for the conditional mean and covariance are obtained [1]

$$\begin{aligned} E(x|z \in A) &= \bar{x} + K [E(z|z \in A) - H\bar{x}] \\ \text{cov}(x|z \in A) &= P + K \text{cov}(z|z \in A) K^T \end{aligned} \quad (6.4)$$

where K is the minimum variance gain matrix, and P is the covariance of x that would be obtained had the measurements not been quantized. These results show that quantization increases the error variance of the estimate as though it was noise added after the measurements.

Suppose a number of measurements are made at different times $t_1, i = 1, 2, \dots$. Since the measurements at different times are correlated, the conditional density of z given the sequence of past measurements will change as more measurements are taken. Thus, even though the equations may be solved recursively, $E(z|z \in A)$ must be recomputed after each measurement; this necessitates a complete solution of the equation after each measurement is taken. We can, however, obtain an approximate recursive estimate by assuming that the conditional distribution of the state just prior to the i^{th} measurement is Gaussian with mean and variance given by Eq. (6.4). For example, let us consider the linear message and observation models

$$x_{k+1} = \phi_k x_k + \Gamma_k w_k$$

$$y_k = c_k x_k$$

$$z_k = y_k + U_k \quad (6.5)$$

where x_k is the $n \times 1$ state vector, and y_k is a $p \times 1$ output vector.

The m -vector w_k is white zero mean gaussian with covariance

$$E\{w_k w_k^T\} = Q_k \delta_{ke}.$$

Let us assume that the observations are quantized so that the estimate of the state x_k must be based on the information that

$$\underline{b}_j \leq z_j < \bar{b}_j \quad j = 1, 2 \dots k \quad (6.6)$$

It can then be shown [2, 3] that the approximate algorithms for estimating x_k for the case when R_k is diagonal are given by

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k D_k \hat{d}_k \quad (6.7)$$

$$\hat{x}_{k|k-1} = \phi_k \hat{x}_{k-1} \quad (6.8)$$

where \hat{d}_k represents a vector whose i^{th} component is

$$[\hat{d}_k]_i = \frac{[f([q_k]_i) - f([u_k]_i)]}{\int_{[q_k]_i}^{[u_k]_i} f([v_k]_i) d[v_k]_i} \quad \begin{aligned} [q_k]_i &= [b_k]_i [H_k x_k]_i \\ [u_k]_i &= [\bar{b}_k]_i [H_k x_k]_i \end{aligned} \quad (6.9)$$

where $f(\cdot)$ denotes the density function of the i^{th} component of v_k , D_k is a diagonal matrix whose elements are

$$[(D_k)^{-1}]_{ii} = \left[[\hat{d}_k]_i [\hat{d}_k]_i - \frac{[q_k]_i f([q_k]_i) - [u_k]_i f([u_k]_i)}{\int_{[q_k]_i}^{[u_k]_i} f([v_k]_i) d[v_k]_i} \right]_{x_k = \hat{x}_{k|k-1}} \quad (6.10)$$

The filter gain and error variance equations are given by

$$K_k = V_k C_k^T [C_k V_k C_k^T + D_k]^{-1} \quad (6.11)$$

$$\begin{aligned} x_k &= \hat{x}_{k|k-1} \\ V_k &= \phi_k [I - K_{k-1} C_{k-1}] V_{k-1} \phi_k^T + \Gamma_k Q_k \Gamma_k^T \end{aligned} \quad (6.12)$$

We now derive the probabilistic estimator for the problem of target tracking described earlier. Thus, the observation model for any sensor, corresponding to the three hypotheses namely (i) noise alone (e.g., cut-off communication link), (ii) false alarm (e.g. thermal or process noise, or clutter from a vehicle not being tracked), and (iii) the vehicle being tracked, is given as follows:

$$H_k^1 : z_k = u_k \quad (6.13a)$$

$$H_k^2 : z_k = c_k \hat{x}(k|k-1) + v_k \quad (6.13b)$$

$$H_k^3 : z_k = c_k x_k + u_k \quad (6.13c)$$

6.2 Probabilistically Supervised Estimation with Quantized Measurements

We use the structure of the probabilistic estimator derived in Chapter 3, in which a set of parallel processing units (PPU's) is used to process the sensor returns. On receiving the quantized observations, each PPU performs a random labelling test to generate a label ℓ_k corresponding to each of the three hypotheses. The label is then accepted as the true hypothesis at that stage. Depending on the label, the observation from the i^{th} sensor is modeled according to Eq. (6.13) by the corresponding PPU. The algorithm for combining the observation models supplied to obtain the track estimate is discussed later.

We can now consider the generation of the random labels.

Suppose there are M sensors and that at stage k , there are n_k returns, $n_k \leq M$. Let ℓ_k^i denote the label generated by the i^{th} PPU at stage k

and let L_k be the vector

$$L_k = [\ell_k^1 \ell_k^2 \dots \ell_k^{n_k}]^T \quad (6.14)$$

Let L_k^M denote the set of all past labels L_j , $j = 1 \dots k$.

Similarly let z_k^i denote the i^{th} sensor return at stage k and

let

$$Z_k = [z_k^1 z_k^2 \dots z_k^{n_k}]^T \quad (6.15)$$

Since the observations are quantized, we can write

$$b_k^i \leq z_k^i < \bar{b}_k^i \quad (6.16)$$

Let B_k denote the hypercube in n_k -dimensional space corresponding to the set of observations at stage k . Then

$$Z_k \in B_k \quad (6.17)$$

Let Z_k^M denote the set of all past observations Z_j , $j = 1, \dots, k$ and

let B_k^M correspond to the quantized observations. Since there are three

hypotheses, the label ℓ_k^i is a ternary random variable. When a new measurement z_{k+1}^i is received, the PPU updates the probabilities associated with the three outcomes on the basis of z_{k+1}^i , the past measurement sequence B_k^M and the past labelling sequence L_k^M . We can then write for these probabilities

$$\begin{aligned} f(\ell_{k+1}^i = H_{k+1}^j | b_{k+1}^i \leq z_{k+1}^i < \bar{b}_{k+1}^i, Z_k^M \in B_k^M, L_k^M) \\ = \frac{\text{Prob}(b_{k+1}^i \leq z_{k+1}^i < \bar{b}_{k+1}^i | \ell_{k+1}^i = H_{k+1}^j, Z_k^M \in B_k^M, L_k^M) f(\ell_{k+1}^i = H_{k+1}^j | Z_k^M \in B_k^M, L_k^M)}{\text{Prob}(b_{k+1}^i \leq z_{k+1}^i < \bar{b}_{k+1}^i | Z_k^M \in B_k^M, L_k^M)} \end{aligned} \quad (6.18)$$

where

$$\text{Prob}(b_{k+1}^i \leq z_{k+1}^i < \bar{b}_{k+1}^i | Z_k^M \in B_k^M, L_k^M)$$

$$= \sum_{j=1}^3 \text{Prob}(b_{k+1}^1 \leq z_{k+1}^1 < \bar{b}_{k+1}^1 | \ell_{k+1}^1 = H_{k+1}^j, Z_k^M \in B_k^M, L_k^M) f(\ell_{k+1}^1 = H_{k+1}^j | Z_k^M \in B_k^M, L_k^M) \\ f(\ell_{k+1}^1 = H_{k+1}^j | Z_k^M \in B_k^M, L_k^M) \quad (6.19)$$

for $i = 1, 2 \dots n_k$; $j = 1, 2, 3$.

If we assume that the hypotheses are independent from stage to stage, we can write

$$f(\ell_{k+1}^1 = H_{k+1}^j | Z_k^M \in B_k^M, L_k^M) = f(\ell_{k+1}^k = H_{k+1}^j) \quad (6.20)$$

It only remains to determine $f(z_{k+1}^k | \ell_{k+1}^1 = H_{k+1}^j, Z_k^M \in B_k^M, L_k^M)$. Let $\hat{x}_{k+1|k}^{(L_k^M)}$ denote the conditional expectation of x_{k+1} conditioned on the sequences of past observations and labels. It then follows from the measurement models of Eq. (6.13) that

$$f(z_{k+1}^1 | \ell_{k+1}^1 = H_{k+1}^j, Z_k^M \in B_k^M, L_k^M) = N_p[0, R_{k+1}] \quad \text{for } j = 1 \quad (6.20a)$$

$$= N_p[C_{k+1} \hat{x}_{k+1|k}^{(L_k^M)}, \Omega_{k+1}^{(L_k^M)}] \quad \text{for } j=2 \quad (6.20b)$$

$$= N_p[C_{k+1} \hat{x}_{k+1|k}^{(L_k^M)}, C_{k+1} V_{k+1|k}^{(L_k^M)} C_{k+1}^T + R_{k+1}] \quad \text{for } j=3 \quad (6.20c)$$

It only remains to evaluate $\hat{x}_{k+1|k}^{(L_k^M)}$. Unfortunately, as discussed earlier, the computation of this estimate is not easy even though the signal and message models are linear, because of the quantized nature of the observations. We can, however, use the approximate algorithms of Sec. 2 to obtain the estimate $\hat{x}_{k+1|k}^{(L_k^M)}$. We will restrict ourselves to the case where R_k is diagonal. The extension to the case of nondiagonal R_k is straight forward. Let C_{k+1}^M denote the matrix

$$C_{k+1}^M = [(C_{k+1}^1)^T | (C_{k+1}^2)^T | \dots | (C_{k+1}^M)^T]^T \quad (6.21)$$

and q_k^1 and u_k^1 denote the vectors

$$q_k^1 = b_k^1 - c_k^1 \hat{x}_{k|k+1} \quad (6.22a)$$

$$u_k^1 = b_k^1 - c_k^1 \hat{x}_{k|k+1} \quad (6.22b)$$

Let d_k^1 denote the vector whose j^{th} component is given by

$$[d_k^1]_j = \frac{f([q_k^1]_j) - f([u_k^1]_j)}{\int_{[q_k^1]_j}^{[u_k^1]_j} f([v_k]_j) d[v_k]_j} \quad (6.23)$$

where $[]_j$ denotes the j^{th} component of a vector and $f(\cdot)$ is the (normal) density function of the j^{th} component of v_k . Let D_k^1 denote the diagonal matrix whose elements are given by

$$[(D_k^1)^{-1}]_{jj} = \left[[d_k^1]_j [d_k^1]_j - \frac{[q_k^1]_j f([q_k^1]_j) - [u_k^1]_j f([u_k^1]_j)}{\int_{[q_k^1]_j}^{[u_k^1]_j} f([v_k]_j) d[v_k]_j} \right] \quad (6.24)$$

Let

$$d_k^M = [(d_k^1)^T; (d_k^2)^T; \dots; (d_k^{n_k})^T]^T \quad (6.25)$$

and

$$D_k^M = \begin{bmatrix} D_k^1 & & \\ & D_k^2 & \\ & & \ddots \\ & & & D_k^{n_k} \end{bmatrix} \quad (6.26)$$

Then the algorithms for determining $\hat{x}_{k+1|k}^{(L_k^M)}$ are given by

$$\hat{x}_{k+1|k} = \phi_k \hat{x}_k \quad (6.27)$$

$$\hat{x}_{k+1} = \hat{x}_{k+1|k} + K_{k+1} D_{k+1}^M d_{k+1}^M \quad (6.28)$$

where the filter gain equation is

$$K_{k+1} = V_{k+1|k} (C_{k+1}^M)^T [C_{k+1}^M V_{k+1|k} (C_{k+1}^M)^T + D_{k+1}^M]^{-1} \quad (6.29)$$

The error variance equation is

$$V_{k+1|k} = \phi_k V_k \phi_k^T + \Gamma_k Q_k \quad (6.30)$$

$$V_{k+1} = [I - K_{k+1} C_{k+1}^M] V_{k+1|k} \quad (6.31)$$

Equations (6.18)-(6.31) constitute a complete set of approximate algorithms for estimating the state of a target in multisensor environments when the observations are in quantized form.

6.3 Simulation

In order to determine the performance of the probabilistically supervised estimator with quantized measurements, the same tracking example considered in earlier chapters was simulated. The system parameters are given as follows

$$x_k = \begin{bmatrix} x_1 - x \text{ position at stage } k \\ x_2 - x \text{ speed at stage } k \\ x_3 - y \text{ position at stage } k \\ x_4 - y \text{ speed at stage } k \end{bmatrix}$$

$$u_k = \begin{bmatrix} x - \text{speed change between stage } k \text{ and } k+1 \\ y - \text{speed change between stage } k \text{ and } k+1 \end{bmatrix}$$

$$\Phi_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$r_k = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad Q_k = \begin{bmatrix} q_{11}(k) & 0 \\ 0 & q_{22}(k) \end{bmatrix}$$

where T is the sampling period.

$$C_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R_k = \begin{bmatrix} r_{11}(k) & 0 \\ 0 & r_{22}(k) \end{bmatrix}$$

In the simulation, the following values are used

$$T = 2 \text{ sec.}$$

$$\rho = 0$$

$$n_{TF} = 0.01$$

x_0 , multivariate normally distributed with dimension four,

$$N_4(0, 0.01)$$

$$Q_k = \text{diag}[0, 0]$$

$$R_k = \text{diag}[0.00828, 0.00828]$$

$$\begin{aligned} \text{Quantizer interval } \Delta &= 2^{-5} \\ &= 0.03125 \end{aligned}$$

$l = 5$, number of sensors

$\text{Prob}(H_k^1) = \text{Prob}(H_k^2) = 0.1$

$\text{Prob}(H_k^3) = 0.8$

A 40 stage process was simulated, and 80 runs were averaged to obtain the mean square error. The mean square error in the x and y positions for the first 10 stages are shown in Fig. 1. Table 1 shows a comparison of the mean square error obtained using the probabilistic estimator as given in [8] with the results obtained with quantized measurements. As can be expected, the error is somewhat larger with quantized measurements.

REFERENCES

- [1] R. E. Curry, Estimation and Control with Quantized Measurements, MIT Press, Cambridge, Mass., 1970.
- [2] K. Clements and R. Haddad, "Approximate Estimation for Systems with Quantized Measurements," IEEE Trans. on Automatic Control, April 1972, pp. 235-239.
- [3] J. F. Metrailler, "Sequential Estimation with Quantized Measurements," Ph.D. Dissertation, Southern Methodist University, Dallas, Texas, April 1972.

TABLE I

MSE Comparison of PSE(Q), PSE
 $P(H_k^1) = P(H_k^2) = 0.1, P(H_k^3) = 0.8$

STAGE	PSE(Q)	PSE[8]
10	.02223	.02071
20	.01122	.00283
30	.00179	.00159
40	.00087	.00065

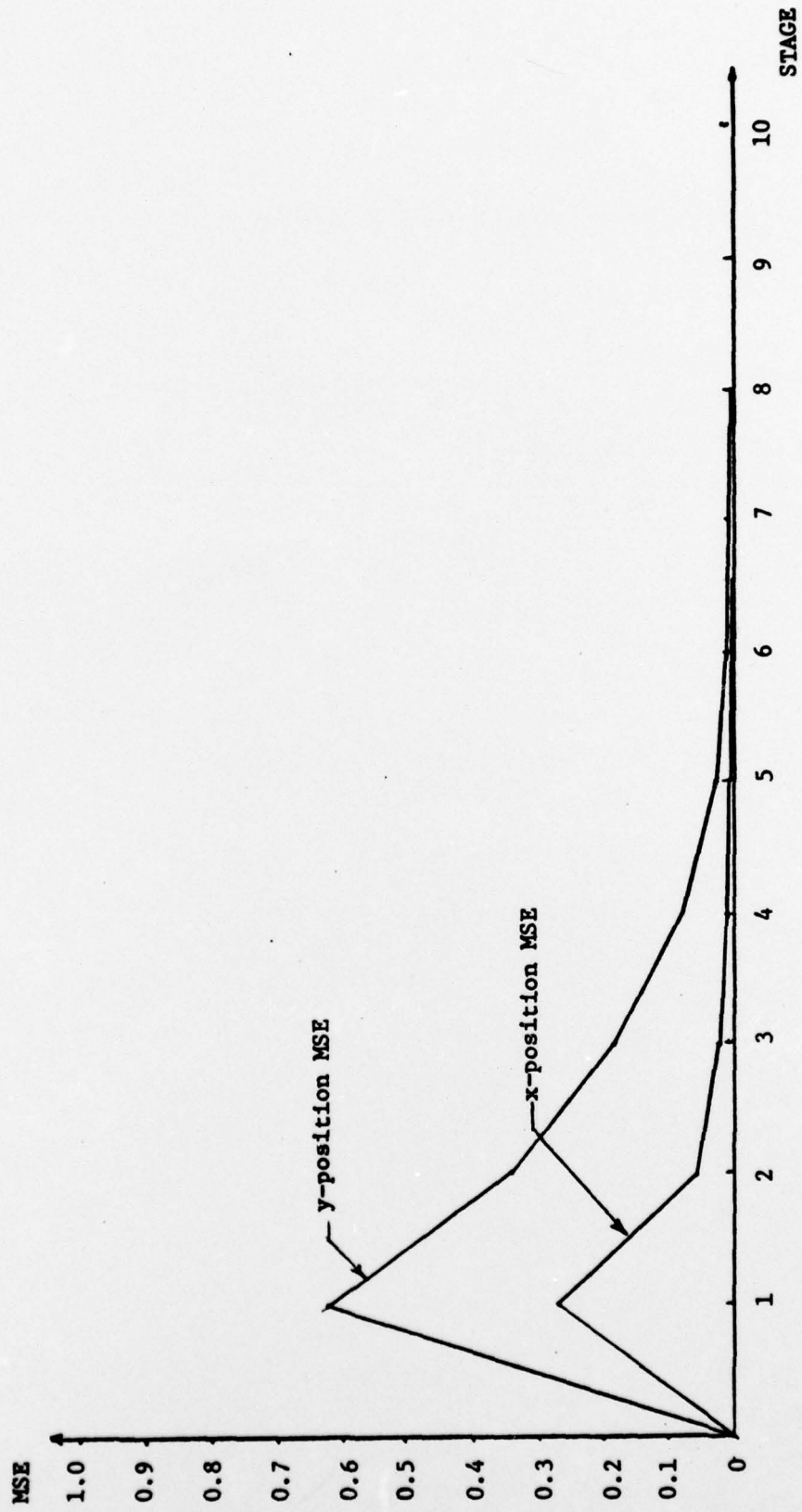


Fig. 1 MSE of PSE with Quantized Measurement

(number of runs - 80)

7. CONCLUSIONS

This research has considered the problem of joint detection and estimation of signals. A specific problem of tracking a target in a multi-target multi-sensor environment was investigated. This problem is somewhat typical of the type of problems encountered in this area. In order to overcome the exploding computational requirements of the problem, a suboptimal scheme which uses a probabilistic judgement to eliminate extraneous returns was derived. Explicit algorithms for tracking the target have been derived. The scheme was compared with two other suboptimal schemes in terms of performance and computational requirements. The tracking algorithms were extended to include cases in which the observations are received over fading media or in quantized form. Simulation results indicate that the performance of the scheme is acceptable under these conditions.

APPENDIX B

Target Tracking over Fading Channels

F. S. CHANG AND M. D. SRINATH

Abstract—The problem of tracking a target in a multitarget environment when the observations are received over a fading channel is considered. The optimal Bayes solution to the tracking problem in such cases involves growing memory and hence is not feasible. A particularly effective suboptimal scheme uses a probabilistic judgment at each stage of the observations to overcome this problem. This concise paper presents an evaluation of the scheme in terms of mean-square error performance when the observations are received over fading channels.

I. INTRODUCTION

Recently there has been an increasing interest in the synthesis of communication systems for space applications. The communication channel modifies the transmitted signal so that the desired message may arrive at the receiver terminal distorted, attenuated, and delayed. The message is usually observed in the presence of additive noise. The distortion introduced by the channel can be characterized by multiplicative

Paper approved by the Associate Editor for Communication Theory of the IEEE Communications Society for publication without oral presentation. Manuscript received June 30, 1975; revised September 8, 1975. This work was supported by AFOSR Grant 74-2727.

The authors are with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275.

Copyright © 1976 by The Institute of Electrical and Electronics Engineers, Inc.
Printed in U.S.A. Annals No. 604C0007

disturbances which are responsible for such phenomena as fading, dispersion, multipath, phase distortion, and time delay.

In recent years considerable effort has been directed toward the application of stochastic estimation theory to the problem of demodulation of modulated signals observed in noise [1]-[4]. Snyder [5] has presented a state-variable model for a wide variety of analog communication systems and channels for a class of Markov processes. He has also presented an analog model for the Rayleigh fading channel in terms of two quadrature multiplicative noise processes which depend upon the frequency of the transmitted signal. Stochastic message and channel disturbance processes were formulated as components of the message model for mathematical convenience. The effect of the modulation process, multiplicative noise, and additive noise was considered in the observation model. The message model was augmented to include the model generating the multiplicative disturbances. A summary of results in this area pertaining to angle modulation is given in [6].

This concise paper presents a scheme for tracking a target in multitarget multisensor environments when the observations presented to the receiver are received over such channels. The dynamic equation of motion of the target being tracked is assumed to be modeled by a set of Gauss-Markov difference equations

$$\begin{aligned} x_{k+1} &= \tilde{\phi}_k x_k + \tilde{\Gamma}_k \omega_k \\ y_k &= h[x_k, k] \end{aligned} \quad (1)$$

where x_k is the $n \times 1$ state vector and y_k is a $p \times 1$ output (message) vector. The m -vector ω_k is white zero-mean Gaussian with covariance $E\{\omega_k \omega_k^T\} = Q_k \delta_{km}$, $\tilde{\phi}_k$ is an $n \times n$ state transition matrix, and $\tilde{\Gamma}_k$ is an $n \times m$ excitation matrix.

If a particular sensor is not locked onto the target being tracked, the corresponding sensor return at any stage can originate from one of three sources (hypotheses), namely, 1) noise alone (e.g., cutoff communication link), 2) false alarm (e.g., thermal or process noise, or clutter from a vehicle not being tracked), and 3) the vehicle being tracked. The receiver performs under uncertainty of reception in a ternary case which considers all the possibilities. A suitable model for the observations under each hypothesis can then be formulated as follows [7].

Under H^1 , the observation z_k is given by

$$H_k^1 : z_k = u_k \quad (2)$$

where the $p \times 1$ vector u_k is a zero-mean white Gaussian sequence with covariance $E\{u_k u_k^T\} = R_k \delta_{km}$. Under hypothesis H^2 , the current observation is not used in track updating and the one-stage prediction is used instead. The effect of the extraneous returns can be modeled as a white zero-mean Gaussian sequence v_k with covariance $E\{v_k v_k^T\} = \Omega_k \delta_{km}$. It is therefore reasonable to model the observation z_k under hypothesis H^2 as

$$H_k^2 : z_k = h[\hat{x}(k | k-1), k] + v_k \quad (3)$$

where $\hat{x}(k | k-1)$ is the one-stage prediction of x_k .

The expression for the covariance kernel Ω_k is given in [9], and for $p = 2$ simplifies to

$$\Omega_k = \frac{C_k V_{k|k-1} C_k^T + R_k}{1 + \frac{\pi}{2} \cdot n_{TF} \cdot \sqrt{1 - \rho^2}} \quad (4)$$

where $V_{k|k-1} = E\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\}$, n_{TF} is the expected number of incorrect returns in one sigma region $\sigma_k = \min\{(C_k V_{k|k-1} C_k^T + R_k)_{ii}\}^{1/2}$, and ρ is the correla-

tion between two measurements, and where

$$C_k = \left(\frac{\partial h}{\partial x_k} \right) x_k = \hat{x}_{k|k-1} \quad (5)$$

The noise sequences ω_k , u_k , and v_k are mutually independent and are uncorrelated with respect to the initial state x_0 , which is multivariate normally distributed of dimension n with mean μ_0 and covariance V_0 , $N_n[\mu_0, V_0]$.

Under H^3 , the observation corresponds to the vehicle being tracked, and hence we have

$$H_k^3 : z_k = h[x_k, k] + u_k \quad (6)$$

It is well known that the optimal Bayes estimator for this problem involves an ever-growing memory and is hence not feasible. This is because the estimate of x_k at each stage depends on the sequence of past true hypotheses. This sequence, however, is unknown, and hence the estimate must be averaged over all possible past sequences. In the pattern recognition context, this is usually referred to as unsupervised learning. (The situation where the sequence of past hypotheses is known corresponds to supervised learning.)

Let Z_k^M denote the set of observations up to stage k , and let α_k^j correspond to a particular sequence of past hypotheses at stage k . We can then write for the estimate

$$\begin{aligned} \hat{x}_k &= E\{x_k | Z_k^M\} \\ &= \sum_{j=1}^{N_k} E\{x_k | Z_k^M, \alpha_k^j\} P(\alpha_k^j | Z_k^M) \end{aligned} \quad (7)$$

where $N_k (=3^k)$ represents the total number of possible sequences at stage k . It is clear that as the number of stages increases, the total number of sequences over which the estimates must be averaged increases very rapidly. We therefore need to resort to suboptimal schemes which use a fixed memory. There are basically two approaches which can be taken to solve this problem. We can use a fixed number, say K , of past observations in arriving at our estimate. Alternately, we can use all past observations in our estimation scheme, but choose only one (or more generally, a fixed number M) of possible past sequences over which to average. A few such schemes have been discussed in [7]-[9].

In this concise paper, we use a probabilistic scheme to decide at each stage which one of the sequences $\alpha_k^1 \dots \alpha_k^{N_k}$ corresponds to the true sequence. The effectiveness of this probabilistically supervised estimator (PSE) is investigated in terms of mean-square error and variance when the observations are received over fading channels.

II. PROBABILISTICALLY SUPERVISED ESTIMATOR

In generating a probabilistically supervised learning sequence, one has to design a probabilistic judgment at each stage of observations on the sensor returns [8], [9]. This judgment is carried out in parallel on each sensor return in such a way that each return is assigned a classification label which then serves as its true hypothesis. We can think of the scheme as consisting of a set of parallel processing units (PPU), one corresponding to each of the sensors, and a central processing unit (CPU) which combines the processed returns from the PPU's to update the estimate of the target track. On receiving the observations, each PPU performs a random labeling test to generate a label l_k corresponding to each of the three hypotheses [8]. The label is then accepted as the true hypothesis at that stage. Depending on the label, the observation from the

i th sensor is modeled according to (2)–(4) by the corresponding PPU. The CPU algorithm for combining the observation models supplied by the PPU's to obtain the track estimate is discussed in a later section.

We can now consider the generation of the random labels. Suppose there are M sensors, and that at stage k there are n_k returns, $n_k \leq M$. Let l_k^i denote the label generated by the i th PPU at stage k , and let $L(k)$ be the vector

$$L(k) = [l_k^1 l_k^2 \dots l_k^{n_k}]^T.$$

Let L_k^M denote the set of all past labels $L(j)$, $j = 1 \dots k$.

Similarly, let z_k^i denote the i th sensor return at stage k , and let

$$Z(k) = [z_k^1 z_k^2 \dots z_k^{n_k}]^T.$$

Let Z_k^M denote the set of all past observations $Z(j)$, $j = 1 \dots k$. Since there are three hypotheses, the label l_k^i is a ternary random variable. When a new measurement z_{k+1}^i is received, the PPU updates the probabilities associated with the three outcomes on the basis of z_{k+1}^i , the past measurement sequence Z_k^M , and the past labeling sequence L_k^M . We can then write for these probabilities

$$f(l_{k+1}^i = H_{k+1}^j | z_{k+1}^i, Z_k^M, L_k^M) = \frac{f(z_{k+1}^i | l_{k+1}^i = H_{k+1}^j, Z_k^M, L_k^M) \cdot f(l_{k+1}^i = H_{k+1}^j | Z_k^M, L_k^M)}{f(z_{k+1}^i | Z_k^M, L_k^M)} \quad (8)$$

where

$$\begin{aligned} f(z_{k+1}^i | Z_k^M, L_k^M) &= \sum_{j=1}^3 f(z_{k+1}^i | l_{k+1}^i = H_{k+1}^j, Z_k^M, L_k^M) \\ &\quad \cdot f(l_{k+1}^i | Z_k^M, L_k^M), \\ &\text{for } i = 1, 2, \dots, n_k; j = 1, 2, 3. \end{aligned} \quad (9)$$

If we assume that the hypotheses are independent from stage to stage, we can write

$$f(l_{k+1}^i = H_{k+1}^j | Z_k^M, L_k^M) = f(l_{k+1}^i = H_{k+1}^j). \quad (10)$$

It only remains to determine $f(z_{k+1}^i | l_{k+1}^i = H_{k+1}^j, Z_k^M, L_k^M)$. Unfortunately, the determination of this density function is extremely difficult because of the nonlinear nature of the observation models. We can seek to approximate this density function by a Gaussian density which has the same first two moments. We then obtain

$$f(z_{k+1}^i | l_{k+1}^i = H_{k+1}^j, Z_k^M, L_k^M) = N_p[0, R_{k+1}^i], \quad \text{for } j = 1 \quad (11)$$

$$= N_p[h^i(\hat{x}_{k+1|k}(L_k^M), k+1), \Omega_{k+1}^i(L_k^M)], \quad \text{for } j = 2 \quad (12)$$

$$\begin{aligned} &= N_p[h^i(\hat{x}_{k+1|k}(L_k^M), k+1), \\ &\quad C_{k+1}^i V_{k+1|k}(L_k^M)(C_{k+1}^i)^T + R_{k+1}^i], \\ &\quad \text{for } j = 3 \end{aligned} \quad (13)$$

where $N_p(m, \sigma^2)$ denotes the p -variate Gaussian density function with mean m and variance σ^2 , and R_k , Ω_k , and C_k are as discussed earlier. The superscript i denotes that the quantities correspond to the i th sensor.

III. AUGMENTED MODELS

Typically, the signal nonlinearly modulates a sinusoidal carrier that arrives at the receiver distorted by disturbance processes encountered in the transmission media. These disturbance processes usually include both additive noise and multiplicative noise. They interact with the channel inputs in a randomly time-varying fashion. In particular, the multiplicative Gaussian channel noise sequences b_k and c_k are assumed to be generated by linear time-invariant models described by two one-dimensional first-order discrete stochastic equations:

$$b_{k+1} = \phi_k^b b_k + [1 - \phi_k^b] m_k^b + \omega_k^b \quad (14)$$

and

$$c_{k+1} = \phi_k^c c_k + [1 - \phi_k^c] m_k^c + \omega_k^c \quad (15)$$

where ω_k^b and ω_k^c are statistically independent scalar zero-mean white Gaussian sequences with variances q_k^b and q_k^c , respectively, and m_k^b and m_k^c represent the time-invariant means for the multiplicative sequences b_k and c_k , respectively. Furthermore,

$$m_k^b = 0, \quad \text{for Rayleigh channel}$$

$$m_k^b = \gamma, \quad \text{for Rician channel}$$

and γ is the specular component [9].

In order to complete the augmented message model, one should adjoin the message equation of (1) and (14) and (15) in any order. For instance, one may write

$$\bar{x}_{k+1} = \phi_k \bar{x}_k + \Gamma_k \bar{\omega}_k + \psi_k \bar{m}_k \quad (16)$$

where

$$\begin{aligned} \bar{x}_k &= \begin{bmatrix} x_k \\ b_k \\ c_k \end{bmatrix}, \quad \bar{\omega}_k = \begin{bmatrix} \omega_k \\ \omega_k^b \\ \omega_k^c \end{bmatrix}, \quad \bar{m}_k = \begin{bmatrix} 0 \\ m_k^b \\ m_k^c \end{bmatrix}, \\ \phi_k &= \begin{bmatrix} \bar{\phi}_k & 0 \\ 0 & \phi_k^b & 0 \\ 0 & 0 & \phi_k^c \end{bmatrix}, \quad \Gamma_k = \begin{bmatrix} \bar{\Gamma}_k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

and

$$\psi_k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & (1 - \phi_k^b) & 0 \\ 0 & 0 & (1 - \phi_k^c) \end{bmatrix}. \quad (17)$$

For the augmented state model, the observation model generated by the i th PPU is given by [see (2), (3), and (6)]

$$z_k^i = u_k, \quad \text{if } l_k^i = H_k^1 \quad (18)$$

$$z_k^i = h[\hat{x}_{k|k-1}, k] + v_k, \quad \text{if } l_k^i = H_k^2 \quad (19)$$

and

$$z_k^i = h[\bar{x}_k, k] + u_k, \quad \text{if } l_k^i = H_k^3 \quad (20)$$

where $\hat{x}_{k|k-1}$ represents the one-stage prediction of \bar{x}_{k-1} .

IV. ESTIMATOR ALGORITHMS

We note from (16) and (18)-(20) that the message and observation equations are nonlinear. We again resort to a wide-sense or Gaussian-fit approximation to the relevant *a posteriori* density functions $f(\bar{x}_k | L_k^M, Z_k^M)$ and $f(\bar{x}_{k+1} | L_k^M, Z_k^M)$. Let us make the following definitions:

$$h_{k+1}^M \triangleq \{h^1[\bar{x}_{k+1}, k+1] \cdots h^n[\bar{x}_{k+1}, k+1]\}^T \quad (21)$$

$$\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \triangleq \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1}} \right) \bar{x}_{k+1} = \hat{x}_{k+1|k}(L_k^M) \quad (22)$$

and

$$R_{k+1}^M \triangleq \text{diag} [R_{k+1}^1 \cdots R_{k+1}^n]. \quad (23)$$

Using a first-order Taylor series approximation about the one-stage prediction, the following equations for updating the first two moments of the density functions can easily be obtained. These equations are similar to the extended Kalman filter algorithms [11].

$$\hat{x}_{k+1|k}(L_k^M) = \phi_k \hat{x}_k(L_k^M) + \psi_k u_k \quad (24)$$

$$V_{k+1|k}(L_k^M) = \phi_k V_k(L_k^M) \phi_k^T + \Gamma_k Q_k \Gamma_k^T \quad (25)$$

$$K_{k+1}(L_{k+1}^M) = V_{k+1|k}(L_k^M) \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T \cdot \left[\left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right) \cdot V_{k+1|k}(L_k^M) \cdot \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T + R_{k+1}^M \right]^{-1} \quad (26)$$

$$\hat{x}_{k+1}(L_{k+1}^M) = \hat{x}_{k+1|k}(L_k^M) + K_{k+1}(L_{k+1}^M) [z_{k+1}^M - \hat{h}_{k+1}^M] \quad (27)$$

and

$$V_{k+1}(L_{k+1}^M) = \left[I - K_{k+1}(L_{k+1}^M) \cdot \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T \right] \cdot V_{k+1|k}(L_k^M) \quad (28)$$

where

$$\hat{h}_{k+1}^M = h_{k+1}^M(\hat{x}_{k+1|k}(L_k^M), k+1). \quad (29)$$

Equations (24)-(29), when used in conjunction with the labeling algorithm of (8)-(13), constitute a complete set of algorithms for target tracking. The algorithms are summarized in Table I.

V. SIMULATION RESULTS

We consider the problem of multiple-sensor tracking of an aircraft in the presence of fading. The message model is as in

TABLE I
ALGORITHMS FOR TARGET TRACKING

Augmented Message Model: $\hat{x}_{k+1}^i = \hat{x}_{k+1|k}^i + K_{k+1}^i (z_{k+1}^i - \hat{h}_{k+1}^i)$

$$y_k = h(\bar{x}_k, k) \quad (16)$$

Observation Model: $z_k^i = u_k$ if $l_k^i = H_k^1$

$$z_k^i = h[\bar{x}_{k|k-1}, k] + v_k \text{ if } l_k^i = H_k^2 \quad (18)$$

$$z_k^i = h(\bar{x}_k, k) + u_k \text{ if } l_k^i = H_k^3 \quad (20)$$

$$i = 1, 2, \dots, n_k \leq M.$$

$$\text{Labelling Algorithm: } f(i_{k+1}^1 = H_{k+1}^1, z_{k+1}^1, L_k^M) = \frac{f(z_{k+1}^1 = u_{k+1}^1, z_{k+1}^2, L_k^M) \cdot f(z_{k+1}^1 = H_{k+1}^1, z_{k+1}^2, L_k^M)}{f(z_{k+1}^1 = H_{k+1}^1, z_{k+1}^2, L_k^M)} \quad (8)$$

where

$$f(z_{k+1}^1 = u_{k+1}^1, z_{k+1}^2, L_k^M) = \prod_{j=1}^2 f(z_{k+1}^j = u_{k+1}^j, z_{k+1}^2, L_k^M) f(z_{k+1}^1 = H_{k+1}^1, z_{k+1}^2, L_k^M) \quad (9)$$

$$f(z_{k+1}^1 = H_{k+1}^1, z_{k+1}^2, L_k^M) = f(z_{k+1}^1 = H_{k+1}^1) \quad (10)$$

$$f(z_{k+1}^1 = u_{k+1}^1, z_{k+1}^2, L_k^M) = \begin{cases} N_p[0, R_{k+1}] & \text{if } j = 1 \\ N_p[h(\hat{x}_{k+1|k}(L_k^M), k+1), z_{k+1}^2(L_k^M)] & \text{if } j = 2 \\ N_p[h(\hat{x}_{k+1|k}(L_k^M), k+1), C_{k+1}^1 V_{k+1|k}(L_k^M) (C_{k+1}^1)^T + R_{k+1}^1] & \text{if } j = 3 \end{cases} \quad (11)$$

Estimation Algorithms:

$$\hat{x}_{k+1|k}(L_k^M) = \phi_k \hat{x}_k(L_k^M) + \psi_k u_k \quad (24)$$

$$V_{k+1|k}(L_k^M) = \phi_k V_k(L_k^M) \phi_k^T + \Gamma_k Q_k \Gamma_k^T \quad (25)$$

$$K_{k+1}(L_{k+1}^M) = V_{k+1|k}(L_k^M) \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T \cdot \left[\left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right) \cdot V_{k+1|k}(L_k^M) \cdot \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T + R_{k+1}^M \right]^{-1} \quad (26)$$

$$\hat{x}_{k+1}(L_{k+1}^M) = \hat{x}_{k+1|k}(L_k^M) + K_{k+1}(L_{k+1}^M) [z_{k+1}^M - \hat{h}_{k+1}^M] \quad (27)$$

$$V_{k+1}(L_{k+1}^M) = \left[I - K_{k+1}(L_{k+1}^M) \cdot \left(\frac{\partial h_{k+1}^M}{\partial \bar{x}_{k+1|k}} \right)^T \right] \cdot V_{k+1|k}(L_k^M) \quad (28)$$

[7] in which the parameters of the aircraft being tracked are given as follows:

$$x_k = \begin{bmatrix} x : \text{position at stage } k \\ x : \text{speed at stage } k \\ y : \text{position at stage } k \\ y : \text{speed at stage } k \end{bmatrix}_{(4 \times 1)}$$

$$\omega_k = \begin{bmatrix} \text{change in } x\text{-speed between stage } k \text{ and stage } k+1 \\ \text{change in } y\text{-speed between stage } k \text{ and stage } k+1 \end{bmatrix}_{(2 \times 1)}$$

$$\phi = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}_{(4 \times 4)} \quad \Gamma_k = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}_{(4 \times 2)}$$

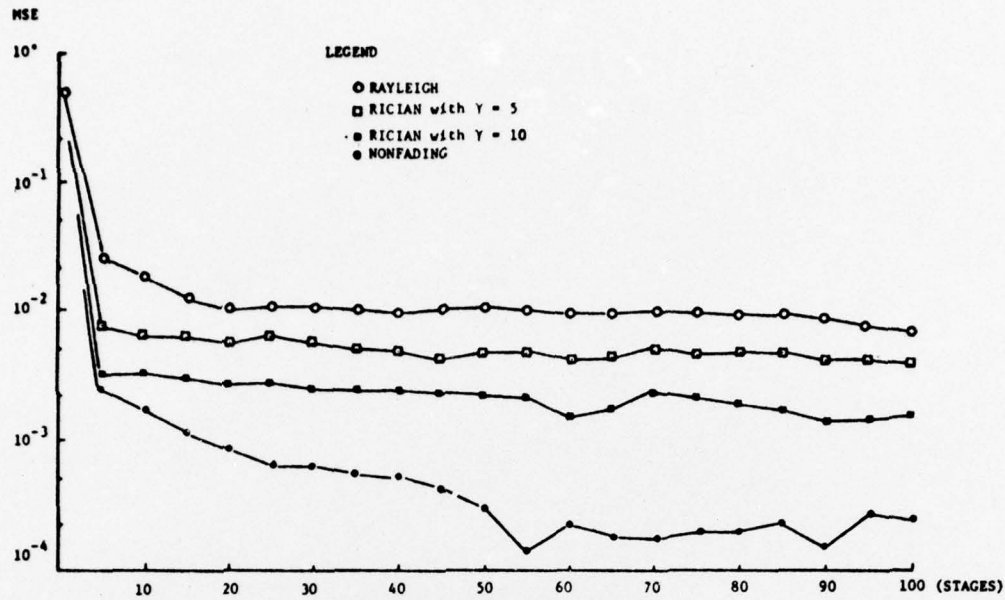


Fig. 1. MSE over fading channels.

and

$$Q_k = \begin{bmatrix} q_{11}(k) & 0 \\ 0 & q_{22}(k) \end{bmatrix}_{(2 \times 2)}$$

where T is the sampling interval.

The additional state space associated with the fading channel processes is represented by two first-order Butterworth spectra given by

$$S_b(\omega) = S_c(\omega) = \frac{2K_b\sigma_b^2}{\omega^2 + K_b^2} \quad (30)$$

where K_b is the half power frequency or one-sided bandwidth and σ_b^2 is the spectral power. The two corresponding one-dimensional discrete equations are

$$b_{k+1} = e^{-K_b T} b_k + [1 - e^{-K_b T}] u_k^b + \omega_k^b \quad (31)$$

and

$$c_{k+1} = e^{-K_b T} c_k + [1 - e^{-K_b T}] u_k^c + \omega_k^c. \quad (32)$$

The excitation covariances are

$$q_k^b = \sigma_b^2(1 - e^{-2K_b T})$$

and

$$q_k^c = \sigma_b^2(1 - e^{-2K_b T}).$$

We assume that the message amplitude modulates a sinusoidal carrier whose frequency is large compared to the significant frequency of the message, so that the signal component of the modulated message has a bandpass spectrum that is essentially disjoint from that of the message. In this case, we can write $h[\bar{x}_k, k]$ in the observation model of (1) as

$$h[\bar{x}_k, k] = [h_1 h_2]^T$$

where

$$h_1 = \sqrt{2}x_1(k)x_5(k)\cos(k\bar{\omega}_1) + \sqrt{2}x_1(k)x_6(k)\sin(k\bar{\omega}_1),$$

$$h_2 = \sqrt{2}x_3(k)x_5(k)\cos(k\bar{\omega}_2) + \sqrt{2}x_3(k)x_6(k)\sin(k\bar{\omega}_2),$$

$\bar{\omega}_1$ and $\bar{\omega}_2$ are the carrier frequencies of x -measurement and y -measurement, respectively, $x_5(k) = b_k$, and $x_6(k) = c_k$. The results for nonfading channels can be obtained by setting $x_5 = 1$ and $x_6 = 0$ in these equations.

The following set of parameters were used in simulations:

$$T = 0.002$$

$$K_b = 30$$

$$P(H_k^1) = P(H_k^2) = 0.1 \quad \text{and} \quad P(H_k^3) = 0.8$$

$$R_k = 0.1656$$

$$\gamma = 5 \quad \text{and} \quad 10$$

$$\bar{\omega}_1 = 31.41 \quad \text{and} \quad \bar{\omega}_2 = 62.82$$

$$S_b(0) = -10 \text{ dB}.$$

A 100-stage process was simulated, and ten runs were averaged to obtain the mean-square errors as well as the variances associated with the first component of the augmented state vector under five ground sensors. The results are shown in Figs. 1 and 2. As may be expected, the curves associated with fading channels are bounded from below by those for the nonfading channel. It may also be expected that the transition from the Rayleigh model to the known signal model may be obtained by varying the value of the specular component γ in the Rician model, with $\gamma = 0$ corresponding to the Rayleigh channel and $\gamma = \infty$ to the completely known signal case. The corresponding normalized mean-square errors are defined as

$$\frac{E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]_{11}}{E[x_k x_k^T]_{11}}$$

(where for any matrix A , $[A]_{ij}$ denotes the ij th entry) and are plotted in Fig. 3.

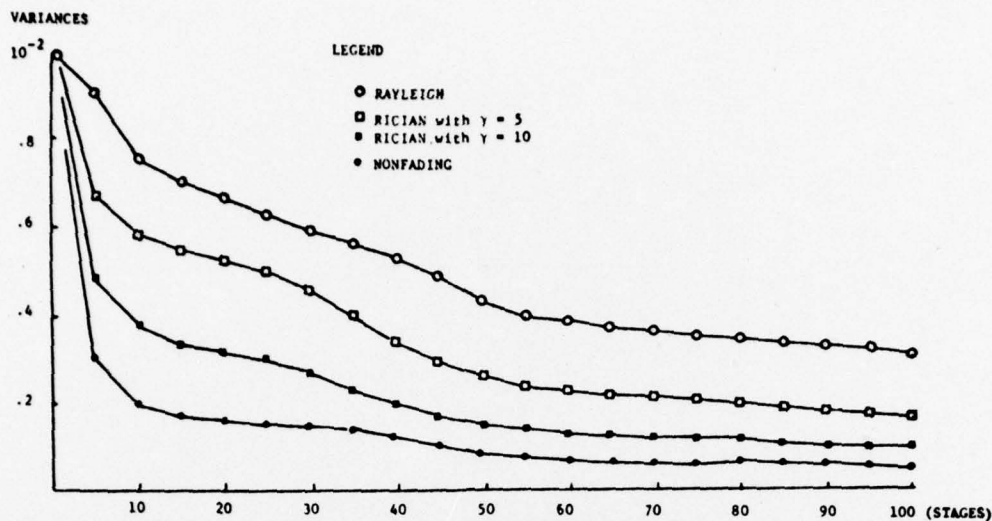


Fig. 2. Plot of variances.

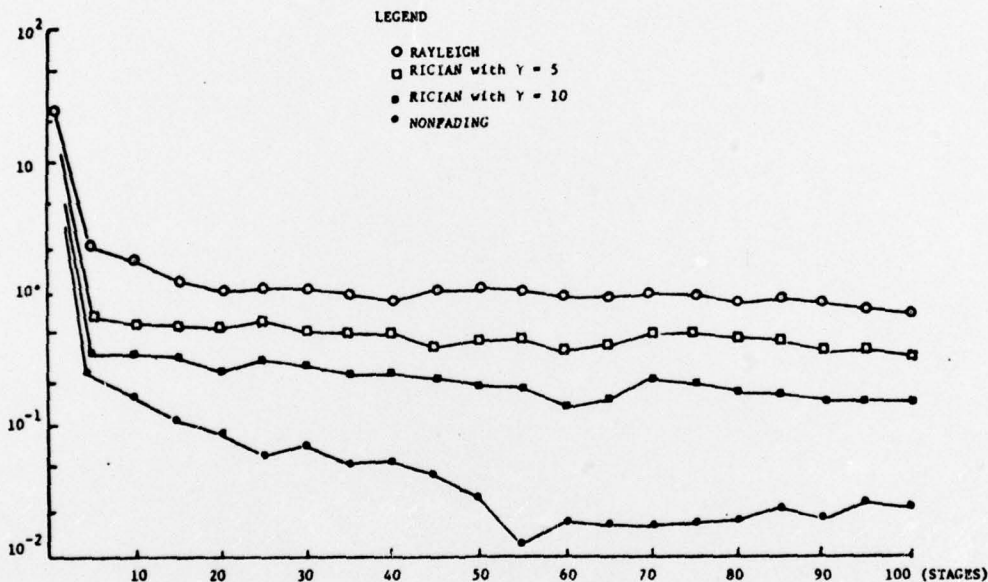


Fig. 3. Normalized MSE.

VI. CONCLUSIONS

This concise paper has presented a set of algorithms for tracking a target in a multitarget environment when the observations are received over a fading channel. The problem of exploding memory is overcome by using a probabilistic scheme in which a random label is accepted as the true hypothesis. The effects of the fading channel are taken into account by augmenting the system model suitably. Since the observation model is nonlinear, the computation of the relevant density functions, both for obtaining the random label and for state estimation, is difficult. An approximate scheme which uses a Gaussian fit to these density functions is used to obtain explicit algorithms. The simulations show that the scheme gives acceptable performance.

REFERENCES

- [1] H. L. Van Trees, *Detection, Estimation, and Modulation, Part II: Nonlinear Modulation Theory*, John Wiley and Sons, Inc., New York, 1971.
- [2] C. N. Kelly and S. C. Gupta, "Digital Time Demodulation of Continuous Time Signals," *IEEE Trans. Information Theory*, Vol. IT-18, pp. 488-493, July 1972.
- [3] J. H. Painter and S. C. Gupta, "Recursive Ideal Observer Detection of Known M -ary Signals in Multiplicative and Additive Gaussian Noise," *IEEE Trans. Communications*, Vol. COM-21, pp. 948-953, Aug. 1973.
- [4] D. L. Snyder, "The State-Variable Approach in Analog Communication Theory," *IEEE Trans. Information Theory*, Vol. IT-14, pp. 94-104, Jan. 1968.
- [5] —, *The State-Variable Approach to Continuous Estimation*, Research Monograph No. 51, The M.I.T. Press, Cambridge, Massachusetts, 1969.
- [6] M. T. Dharamsi and S. C. Gupta, "Discrete-Time Demodulation of Angle-Modulated Analog Signals over Fading Channels," *Information Science*, 1975.
- [7] R. A. Singer, R. G. Sea, and K. B. Housewright, "Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multi-target Environments," *IEEE Trans. Information Theory*, Vol. IT-20, pp. 423-432, July 1974.
- [8] A. K. Agrawala, "Learning with a Probabilistic Teacher," *IEEE Trans. Information Theory*, Vol. IT-16, pp. 373-379, July 1970.

- [9] F. S. Chang, "Estimation and Detection in Uncertain Environments," Ph.D. Dissertation, Southern Methodist University, May 1975.
- [10] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
- [11] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill, New York, 1971.

APPENDIX A

STATISTICS OF EXTRANEIOUS RETURNS

The output equation for extraneous sensor returns (hypothesis H_2) was modeled as

$$z_k = C_k \hat{x}_{k|k-1} + v_k \quad (A.1)$$

where v_k is assumed to be Gaussian, white with zero mean and covariance Ω_k . The assumption of whiteness of the miscorrelations has been discussed in [1] where it has been shown by Monte Carlo simulations that this is a good representation for predicting performance in dense environments.

In order to derive an expression for Ω_k , we can follow the derivation in [2] and assume all returns fall within a region (gate) G in the observation space and let the size of the gate tend to infinity. In order for (A.1) to represent an incorrect return, the corresponding return must be the closest to $C_k \hat{x}_{k|k-1}$ given that the closest point is an incorrect return. Let

$$\psi_k = C_k V_{k|k-1} C_k^T + R_k \quad (A.2)$$

where $V_{k|k-1}$ and R_k have been defined in Chapter 2 and let

$$\sigma_k = \min [(\psi_k)_{ii}]^{1/2} \quad (A.3)$$

and let M be some integer.

We then define a region \tilde{G} as

$$\tilde{G} = \{\tilde{z}_k | \tilde{z}_k = z_k - C_k \hat{x}_{k|k-1}; ||z_k||^2_{\psi_k} - 1 \leq M\sigma_k\} \quad (A.4)$$

Similarly let \tilde{D} be the region such that

$$\tilde{D} = \{\tilde{y}_k | \tilde{y}_k = C_k \tilde{x}_{k|k-1} + v_k; ||y_k||^2_{\psi_k} - 1 \leq ||v_k||^2_{\psi_k} - 1\} \quad (A.5)$$

It then follows that there should be no other returns with residuals in the region $\tilde{G} \cap D$, so that the density function for v_k given that v_k is an extraneous return is

$$\begin{aligned}
& f(v_k | v_k \text{ is residual of an incorrectly correlated return}) \\
& = f(v_k | \text{no other return in } \tilde{G} \cap \tilde{D}) \\
& = \frac{P[\text{no other return in } \tilde{G} \cap \tilde{D} | v_k] f(v_k)}{P[\text{no other return in } \tilde{G} \cap \tilde{D}]} \quad (A.6)
\end{aligned}$$

If we assume that all incorrect returns are uniformly distributed for $v_k \in \tilde{G}$ and that $\tilde{G} \cap \tilde{D}$ can be replaced by \tilde{G} with little error (see [2]), we can write

$$\begin{aligned}
& P[\text{no other return in } \tilde{G} \cap \tilde{D} | v_k] \\
& = P[\text{no correlation is made} \mid \text{proper return is detected}] P_D \\
& + P[\text{no correlation is made} \mid \text{proper return is not detected}] (1 - P_D) \\
& = \left[1 - \frac{P_D}{(2\pi)^{n/2} |\psi_k|^{1/2}} \int_{\tilde{G}} \exp\left\{-\frac{1}{2} \left| \frac{\tilde{y}_k}{\psi_k} \right|^2 - 1\right\} d\tilde{y}_k \right] \\
& \cdot \exp\left\{-\frac{1}{4} \frac{n_{TF}}{|\psi_k|^{1/2}} \int_{\tilde{D}} d\tilde{y}_k\right\} \quad (A.7)
\end{aligned}$$

where P_D denotes the probability of detection. We note that the denominator of Eq. (A.6) is just a normalizing constant. The preceding equation for $p = 2$ reduces to [2]

$$\begin{aligned}
& f(v_k | \text{no other return in } \tilde{G} \cap \tilde{D}) \\
& = \text{const.} \exp\left\{-\frac{1}{2} \left| \frac{v_k}{\Omega_k} \right|^2 - 1\right\} \quad (A.8)
\end{aligned}$$

where

$$\Omega_k = \frac{\psi_k}{1 + \frac{\pi}{2} \cdot n_{TF} \sqrt{1 - \rho^2}}$$

Here ρ is the correlation between two measurements, and n_{TF} is the expected number of incorrect returns in a one- σ_k region.

REFERENCES

1. R. A. Singer, R. G. Sea and K. B. Housewright. "Comparison of Theoretical and Simulated Performance of Optimal and Suboptimal Filters in a Dense Multi-target Environment", Proc. IEEE Conf. on Decision and Control, New Orleans, La, Dec. 1972, pp 700-701.
2. R. G. Sea, "An Efficient Suboptimal Decision Procedure for Associating Sensor Data with Stored Tracks in Real-Time Surveillance Systems" Proc. IEEE Decision and Control Conf., Miami Beach, Fla. Dec. 1971, pp 33-37.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
18 AFOSR - TR - 76 - 1119 /			
4. TITLE (and Subtitle)		5. TYPE OF REPORT & PERIOD COVERED	
6 DETECTION AND ESTIMATION OF SIGNALS IN UNCERTAIN ENVIRONMENTS.		9 Final rept.	
7. AUTHOR(s)		6. PERFORMING ORG. REPORT NUMBER	
10 M. D. Srinath	12 77p.	15 ✓ AF- AFOSR - 2727 - 74	
9. PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
Southern Methodist University Department of Electrical Engineering Dallas, Texas 75275		16 AF-2304A 12 5 61102F 2304/A5	
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE	
Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		11 31 Aug 1976	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES	
		72	
		15. SECURITY CLASS. (of this report)	
		UNCLASSIFIED	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Detection, estimation, target tracking, probabilistic learning, exploding memory, optimal Bayesian learning, uncertain observation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In many communication problems, it becomes necessary to simultaneously detect and extract the signal so that these two operations must be jointly optimized in order to improve receiver performance. The receiver structures so obtained are considerably more complicated than conventional receivers and it is not clear that this added complexity results in a significant improvement in receiver performance. This report investigates a specific suboptimal scheme for estimating a signal in the presence of uncertainty regarding its reception. The scheme uses a probabilistic judgement to determine the presence of the

DD FORM 1473 1 JAN 73

EDITION OF 1 NOV 68 IS OBSOLETE

408 600 UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20 Abstract

signal in the observations. In order to obtain explicit estimation algorithms and to facilitate performance evaluation, a specific problem of signal estimation under uncertainty, namely the tracking of a target in a multi-target multi-sensor environment is considered. Explicit estimation algorithms are developed and the performance of the receiver is evaluated under conditions involving fading media or quantized observations.